

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2004, 14:00:34 ; Search time 194 Seconds  
(without alignments)  
513.092 Million cell updates/sec

Title: US-09-820-003C-2  
Perfect score: 886  
Sequence: 1 MMSMNPEYDYLFLKLLIGDS.....EKSNVKIQSTPVRKSGGCC 173

Scoring table: BIOSOM62  
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt 02:.\*  
1: uniprot\_sprot:.\*  
2: uniprot\_trembl:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	860	97.1	205	1	RB1A CANFA
2	860	97.1	205	1	RB1A_HUMAN
3	860	97.1	205	1	RB1A_MOUSE
4	860	97.1	205	1	RB1A_RAT
5	860	97.1	205	2	AAH66662
6	860	97.1	205	2	CAE11872
7	860	97.1	205	2	AAH62594
8	860	97.1	205	2	BAC28697
9	853	96.3	253	2	Q6ZPF0
10	853	96.3	253	2	BAC98287
11	831.5	93.8	204	2	O6GLH0
12	831.5	93.8	204	2	Q7ZXF7
13	828	90.3	202	1	RB1A_DISOM
14	800.5	90.3	201	2	Q7ZS20
15	800.5	90.3	201	2	AAH62857
16	799.5	90.2	201	2	Q7ZUV5
17	766.5	86.5	201	1	RB1B_HUMAN
18	763.5	86.2	201	1	O6F1G4
19	760.5	85.8	201	1	RB1B_RAT
20	759.5	85.7	201	1	RB1B_MOUSE
21	752	84.9	301	2	Q7LPR7
22	733	82.7	205	1	RB1A_LYMST
23	683	77.1	205	2	O18133
24	671	75.7	202	2	O16143
25	670	75.6	164	2	O610G4
26	670	75.6	164	2	AAH71442
27	660	74.5	205	2	Q9UAC6
28	658.5	74.3	203	1	YPT1_NEUCR
29	655.5	74.0	201	2	Q9HET3
30	653.5	73.8	201	2	Q9P8J7
31	646.5	73.0	203	1	YPT1_CHLRE

32	646.5	73.0	203	1	YPT1_SCHPO	P1620 schizosacch
33	644	72.7	202	2	Q40203	Q40203 lotus japon
34	642.5	72.5	203	1	YPT1_VOLCA	P1584 volvox cart
35	642	72.5	202	2	Q9SEH3	Q9SEH3 arabidopsis
36	640.5	72.3	212	2	Q7Q874	Q7Q874 anophelis g
37	640	72.2	202	2	Q9HDT5	Q9HDT5 trichoderma
38	640	72.2	202	2	O6C9V1	O6C9V1 yarrowia li
39	636.5	71.8	203	2	Q08154	Q08154 plasm saliv
40	635	71.8	243	2	Q7RVH8	Q7RVH8 neurospora
41	634.5	71.6	202	2	Q8K4S8	Q8K4S8 arabidopsis
42	633	71.4	202	2	Q9FPV4	Q9FPV4 arabidopsis
43	632.5	71.4	203	2	O24112	O24112 nicotiana p
44	632	71.3	202	2	Q9SXT5	Q9SXT5 cicier arlet
45	631.5	71.3	201	1	YPT1_PHYIN	O01890 phycophthor

## ALIGNMENTS

RESULT 1  
RB1A\_CANFA STANDARD; PRT; 205 AA.  
ID RB1A\_CANFA  
AC P62822; P05711;  
DT 01-NOV-1988 (Rel. 09, Created)  
DT 01-OCT-2004 (Rel. 45, Last sequence update)  
DT 01-OCT-2004 (Rel. 45, Last annotation update)  
DE Ras-related protein Rab-1A.  
GN Name=RB1A; Synonyms=RAB1;  
OS Canis familiaris (Dog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
OX NCBI\_TaxID=9615;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Cocker spaniel;  
RX MEDLINE=91061765; PubMed=2123294;  
RA Chavrier P., Vingron M., Sander C., Simons K., Zerial M.;  
RT "Molecular cloning of YPT1/SEC4-related cDNAs from an epithelial cell  
line.";  
RL Mol. Cell. Biol. 10:6578-6585 (1990).  
RN [2]  
RP ISOPRENOLD.  
RX MEDLINE=91296801; PubMed=1648736;  
RA Khosravi-Far R., Lutz R.J., Cox A.D., Conroy L., Bourne J.R.,  
RA Sinensky M., Balch W.E., Buss J.B., Der C.J.;  
RT "Isoprenoid modification of Rab proteins terminating in CC or CXC  
motifs.";  
RL Proc. Natl. Acad. Sci. U.S.A. 88:6264-6268 (1991).  
CC -1- FUNCTION: Probably required for transit of protein from the ER  
through Golgi compartment. Binds GTP and GDP and possesses  
intrinsic GTPase activity.  
CC -1- SUBCELLULAR LOCATION: Golgi.  
CC -1- PTM: Phosphorylated by CDC2 kinase during mitosis (By similarity).  
CC -1- SIMILARITY: Belongs to the small GTPase superfamily. Rab family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC -----  
DR EMBL; X56384; CAB56775.1; -.  
DR HSSP; P07560; 1G16.  
DR InterPro; IPR001806; Ras trnsfmg.  
DR InterPro; IPR005225; Small\_GTP.  
DR Pfam; PF00071; Rasg\_1.  
DR PRINTS; PR00449; RASTRNSFRMG.  
DR TIGRFAMs; TIGR00231; small\_gtp\_1.  
DR Endoplasmic reticulum; Golgi stack; GTP-binding; Lipoprotein;  
KW Prenylation; Protein transport.  
FT NP\_BIND 18 25 GTP (By similarity).

FT NP\_BIND 66 70 GTP (By similarity).  
 FT NP\_BIND 124 127 GTP (By similarity).  
 FT DOMAIN 40 48 Effector region (By similarity).  
 FT LIPID 204 204 S-geranylgeranyl cysteine.  
 FT LIPID 205 205 S-geranylgeranyl cysteine.  
 SQ SEQUENCE 205 AA; 22678 MW; B2AF4B3B0FB17D6 CRC64;

Query Match 97.1%; Score 860; DB 1; Length 205;  
 Best Local Similarity 84.4%; Pred. No. 2e-62;  
 Matches 173; Conservative 0; Mismatches 0; Indels 32; Gaps 1;

QY 1 MSSNNPEVDYLFKLLIGDSGVGKSCLLRFADPTTSTSYSTISITGVDPKIRITIELDGKTI 60  
 DB 1 MSSNPEVDYLFKLLIGDSGVGKSCLLRFADPTTSTSYSTISITGVDPKIRITIELDGKTI 60  
 QY 61 KLIQI-----ESPNNVKQMLQSIDRYASENVKL 88  
 DB 61 KLIQIDPTAGQERFRTTSSYRGAGHIIIVYDVTQESFNNVKQMLQSIDRYASENVKL 120  
 QY 89 LVGNKCDLTTKKVVDYTTAKFADSLGIPILETSAKNNTVQSPFMTAAEIKKRMGPGA 148  
 DB 121 LVGNKCDLTTKKVVDYTTAKFADSLGIPILETSAKNNTVQSPFMTAAEIKKRMGPGA 180  
 QY 149 TAGGAEKSNVKIQSTPPVQSGGGGCC 173  
 DB 181 TAGGAEKSNVKIQSTPPVQSGGGGCC 205

RESULT 2  
 RBLA HUMAN  
 ID RBLA HUMAN STANDARD; PRT; 205 AA.  
 AC P62820; P11476; Q96N61; Q9Y3T2;  
 DT 01-OCT-1989 (Rel. 12, Created)  
 DT 01-OCT-1989 (Rel. 12, Last sequence update)  
 DT 01-OCT-2004 (Rel. 45, Last annotation update)  
 DE Ras-related protein Rab-1a (YPT1-related protein).  
 GN Name=RAB1A; Synonyms=RAB1;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORN 1).  
 RX MEDLINE=89308666; PubMed=2501306;  
 RA Zahracou A., Touchot N., Chardin P., Tavittian A.;  
 RT "The human Rab genes encode a family of GTP-binding proteins related  
 RT to yeast YPT1 and SEC4 products involved in secretion.";  
 RL J. Biol. Chem. 264:12394-12401 (1989).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORN 3).  
 RC TISSUE=Brain;  
 RX MEDLINE=21154917; PubMed=11230166; DOI=10.1101/gr.154701;  
 RA Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glaesl S.,  
 RA Ansorge W., Boecker M., Bioecker H., Bauersachs S., Blum H.,  
 RA Laber J., Duestenhoef A., Beyer A., Koehner K., Strack N.,  
 RA Mewes H.-W., Oettermann B., Obermaier B., Tampe J., Heubner D.,  
 RA Wandt R., Korn B., Klein M., Poustka A.;  
 RT "Towards a catalog of human genes and proteins: sequencing and  
 RT analysis of 500 novel complete protein coding human cDNAs";  
 RL Genome Res. 11:422-435 (2001).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORN 1).  
 RC TISSUE=Colon;  
 RA Koehner K., Beyer A., Mewes H.-W., Weil B., Amid C., Oeanger A.,  
 RA Podo G., Han M., Wiemann S.;  
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A. (ISOFORN 2).  
 RX PubMed=14702039; DOI=10.1038/ng1285;  
 RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,  
 RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,  
 RA Sekine M., Ohyashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,  
 RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,

RA Nagahari K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,  
 RA Shitatori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,  
 RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikawa E.,  
 RA Omura Y., Abe K., Kamihara K., Katuta N., Sato K., Tanikawa M.,  
 RA Yamazaki M., Nimomiyama K., Iehibashi T., Yamashita H., Murakawa K.,  
 RA Fujimori K., Tani H., Kimata M., Watanabe M., Hirooka S., Chiba Y.,  
 RA Ishida S., Ono Y., Takiguchi S., Watanabe S., Yoshida M., Houta T.,  
 RA Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,  
 RA Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K., Arita M.,  
 RA Imose N., Miasashiro K., Yuki H., Oshima A., Sasaki N., Aotsuka S.,  
 RA Yoshikawa Y., Matsunawa H., Ichihara T., Shibata N., Sano S.,  
 RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,  
 RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,  
 RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,  
 RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,  
 RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujitawa T.,  
 RA Ono T., Yamada K., Fujii Y., Ozaki K., Hiroo M., Omori Y.,  
 RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,  
 RA Ohtani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,  
 RA Matsunura K., Nakajima Y., Mizuno T., Moritaga M., Sasaki M.,  
 RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,  
 RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,  
 RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuno Y., Yamashita R.,  
 RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;  
 RT "Complete sequencing and characterization of 21,243 full-length human  
 RT cDNAs.";  
 RL Nat. Genet. 36:40-45 (2004).  
 RN [5]  
 RP SEQUENCE FROM N.A. (ISOFORN 1).  
 RC TISSUE=Brain;  
 RA Publ H.L. III, Ikeda S.R., Aronstam R.S.;  
 RT "cDNA clones of human proteins involved in signal transduction  
 RT sequenced by the Guttrich cDNA resource center (www.cdna.org).";  
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A. (ISOFORN 1).  
 RC TISSUE=Brain, and Placenta;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.2426038999;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L.H., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Datchenko L., Muzyk D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richardson D.K., Muzny K.C., Hale S., Garcia A.M., Gay L.J., Hultky S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman D.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butlerfield Y.S.N., Krzyzanski M.I., Skalska U., Smalins D.E.,  
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [7]  
 RP PHOSPHORYLATION BY CDC2  
 RX MEDLINE=91218852; PubMed=1902553;  
 RA Bailly E., Mc Caffrey M., Touchot N., Zahracou A., Goud B., Bornens M.;  
 RT "Phosphorylation of two small GTP-binding proteins of the Rab family  
 RT by p34cdc2.";  
 RL Nature 350:715-718 (1991).  
 RN [8]  
 RP FUNCTION: Probably required for transit of protein from the ER  
 RP through Golgi compartment. Binds GTP and GDP and possesses  
 RP intrinsic GTPase activity.  
 RP SUBCELLULAR LOCATION: Golgi.  
 RP -1- ALTERNATIVE PRODUCTS:  
 RP Event=Alternative splicing; Named isoforms=3;  
 RP Name=1;  
 RP IsoId=P62820-1, P11476-1; Sequence=Displayed;

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2004, 14:08:40 ; Search time 38 seconds  
(without alignments)  
438.039 Million cell updates/sec

Title: US-09-820-003C-2  
Perfect score: 886  
Sequence: 1 MSSMNPEDYLFKLLIGDS.....EKNVKTQSTPVKSGGGCC 173

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR\_79:.\*  
2: PIR1:.\*  
3: PIR2:.\*  
4: PIR3:.\*  
5: PIR4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	860	97.1	205	1	TVDGRP
2	860	97.1	205	1	TVHUP
3	860	97.1	205	1	TVMSYP
4	849	95.8	205	1	TVRTYP
5	812.5	91.7	201	2	D38625
6	771.5	87.1	201	2	S06147
7	733	82.7	205	2	S38339
8	671	75.7	202	2	JB0318
9	660	74.5	203	2	S30096
10	658.5	74.3	203	2	S30096
11	646.5	73.0	203	2	YC105
12	646.5	73.0	203	2	YC105
13	642.5	72.5	203	2	YC1247
14	631.5	71.3	201	2	UC5337
15	627.5	70.8	203	2	S34253
16	623.5	70.4	206	2	S04590
17	619.5	69.9	258	2	B86153
18	619	69.9	202	2	S41430
19	618.5	69.8	203	2	B38202
20	613.5	69.2	202	2	S38740
21	612	69.1	221	2	H71444
22	598	67.5	218	2	T07609
23	590.5	66.3	202	2	S72515
24	587	66.3	201	2	S39565
25	574	64.8	206	2	T14391
26	572.5	64.6	196	2	PS0279
27	571	64.4	206	1	TVBYQ2
28	523	59.0	208	2	A38202
29	468.5	52.9	208	2	A34716

30	461	52.0	203	2	B34716	GTP-binding protei
31	447	50.5	216	2	T48378	GTP-binding protei
32	442	49.9	215	2	T15655	GTP-binding protei
33	439.5	49.6	215	2	S57478	GTP-binding protei
34	438.5	49.5	216	2	US0640	GTP-binding protei
35	435.5	49.2	200	2	S12790	GTP-binding protei
36	430.5	48.6	215	2	S57462	GTP-binding protei
37	429.5	48.5	216	2	S57471	GTP-binding protei
38	429.5	48.5	222	2	T14405	GTP-binding protei
39	428.5	48.4	216	2	T45901	GTPase ARAB8 - Ar
40	426.5	48.1	215	2	S57474	GTP-binding protei
41	425.5	48.0	216	2	S33900	GTP-binding protei
42	418	47.2	207	2	B49647	GTP-binding protei
43	418	47.2	207	2	B36364	GTP-binding protei
44	410.5	46.3	206	2	T78851	GTP-binding protei
45	409.5	46.2	209	2	B38625	GTP-binding protei

#### ALIGNMENTS

##### RESULT 1

TVDGRP GTP-binding protein Rab1 - dog

C:Species: Canis lupus familiaris (dog)

C:Date: 19-Feb-1994 #sequence revision 06-Dec-1996 #text\_change 19-Jan-2001

C:Accession: S19104; A36364; S15600

R:Zerial, M. Submitted to the EMBL Data Library, August 1990

A:Reference number: S19104

A:Accession: S19104

A:Molecule type: mRNA

A:Residues: 1-205 <ZER>

A:Cross-references: EMBL:X56384

R:Chavrier, P.; Vingron, M.; Sander, C.; Simons, K.; Zerial, M.

Mol. Cell. Biol. 10, 6578-6585, 1990

A:Title: Molecular cloning of YPT1/SEC4-related cDNAs from an epithelial cell line.

A:Reference number: A36364; MUID:91061765; PMID:2123294

A:Accession: A36364

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-157, 'EK', 160-205 <CHA>

A:Cross-references: GB:X56384; NID:9913

C:Function:

A:Description: probably involved in protein transport from the endoplasmic reticulum thr

C:Superfamily: ras transforming protein; translation elongation factor Tu homology

C:Keywords: endoplasmic reticulum; Golgi apparatus; GTP binding; lipidprotein; membrane p

tein F:12-127/Domain: translation elongation factor Tu homology <FTU>

F:18-25/Region: nucleotide-binding motif A (P-loop)

F:62-67/Region: nucleotide-binding motif B

F:124-127/Region: GTP-binding NKXD motif

F:154-156/Region: GTP-binding SAK/L motif

F:24-25,43,124,125,127,154/Binding site: Mg-GTP (Lys, Ser, Thr, Asn, Lys, Asp, Ser) #stat

F:194/Binding site: phosphate (Ser) (covalent) (by cdc2 kinase) #statue predicted

F:204,205/Binding site: geranyl-geranyl (Cys) (covalent) #statue predicted

Query Match 97.1%; Score 860; DB 1; Length 205;

Best local similarity 84.4%; Pred. No. 1.7e-61;

Matches 173; Conservative 0; Mismatches 0; Indels 32; Gaps 1;

QY	1	MSSMNPEDYLFKLLIGDSGVKSCLLAFADPTYESYITGVDFKRTIELDOKTI	60
DB	1	MSSMNPEDYLFKLLIGDSGVKSCLLAFADPTYESYITGVDFKRTIELDOKTI	60
QY	61	KLOI-----ESFNNVQWLOEIDRYASENVKL	88
DB	61	KLOIWDNAGGERPRTTSSYRGAGHGIIVYDVTQDSFNNVQWLOEIDRYASENVKL	120
QY	89	LVGNKCDLTKKVVDDYTTAKEPADSLGIPLETSKAKNAIVESQFMTAAEIKRNGPGA	148
DB	121	LVGNKCDLTKKVVDDYTTAKEPADSLGIPLETSKAKNAIVESQFMTAAEIKRNGPGA	180

QY 149 TAGGAEKSNVVKIOSTPVKQSGGGCC 173  
 Db 181 TAGGAEKSNVVKIOSTPVKQSGGGCC 205

## RESULT 2

TVHUYF

GTP-binding protein Rab1 - human

N/Alternate names: protein DKFZp564B163.1

C/Species: Homo sapiens (man)

C/Date: 29-Jun-1990 #sequence\_revision 06-Dec-1996 #text\_change 09-Jul-2004

C/Accession: A34323; T08698

R/Zahraoui, A.; Touchot, N.; Chardin, P.; Tavittian, A.

J. Biol. Chem. 264, 12394-12401, 1989

A/Title: The human Rab gene encode a family of GTP-binding proteins related to yeast YF

A/Reference number: A34323; MUID:89308668; PMID:2501306

A/Accession: A34323

A/Molecule type: mRNA

A/Residues: 1-205 &lt;ZAH&gt;

A/Cross-references: UNIPROT:P11476; GB:J04941; GB:M28209; NID:g55059; PIDN:AAA60240.1;

R/Habdruck, R.; Heubner, D.; Wewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, March 1999

A/Reference number: Z16471

A/Accession: T08698

A/Molecule type: mRNA

A/Residues: 1-64,141-205 &lt;MAM&gt;

A/Cross-references: EMBL:AL050268

A/Experimental source: fetal brain; clone DKFZp564B163

C/Genetics:

A/Gene: GDB:RAB1

A/Cross-references: GDB:118857; OMIM:179508

A/Map position: 4P15.31-4P15.31

A/Note: DKFZp564B163.1

C/Function:

A/Description: probably involved in protein transport from the endoplasmic reticulum th

C/Superfamily: ras transforming protein; translation elongation factor Tu homolog

C/Keywords: endoplasmic reticulum; Golgi apparatus; GTP binding; lipoprotein; membrane p

tein

F:1-205/Product: GTP-binding protein Rab1 #status predicted &lt;MAT1&gt;

F:1-64,141-205/Product: GTP-binding protein Rab1, splice variant #status predicted &lt;MAT2

F:12-127/Domains: translation elongation factor Tu homology &lt;ETU&gt;

F:18-127/Region: nucleotide-binding motif A (P-loop)

F:62-67/Region: nucleotide-binding motif B

F:124-127/Region: GTP-binding NXXD motif

F:154-156/Region: GTP-binding SAK/L motif

F:24,25,43,124,125,127,154/Binding site: Mg-GTP (Lys, Ser, Thr, Asn, Lys, Asp, Ser) #sta

F:194/Binding site: phosphate (Ser) (covalent) (by cdcd kinase) #status predicted

F:204,205/Binding site: geranyl-geranyl (Cys) (covalent) #status experimental

Query Match 97.1%; Score 860; DB 1; Length 205;

Best Local Similarity 84.4%; Pred. No. 1.7e-61;

Matches 173; Conservative 0; Mismatches 0; Indels 32; Gaps 1;

QY 1 MSSNNPEYDYLFLKLLIGDSGVKSCLLRFADDTYESYISTIGVDFKIRITELDGKTI 60

Db 1 MSSNNPEYDYLFLKLLIGDSGVKSCLLRFADDTYESYISTIGVDFKIRITELDGKTI 60

QY 61 KLGQI-----ESNNYKQWLOEIDRYASENVNKL 88

Db 61 KLGQIWDTAGQERFRITTSYRGAGHIIIVYDVTQESFNNVQWLOEIDRYASENVNKL 120

QY 89 LVGNKCDLTTKKVVDYTTAKEPADSLGIPLETSAKATNVQSPFMTMAAEIKKMGPGA 148

Db 121 LVGNKCDLTTKKVVDYTTAKEPADSLGIPLETSAKATNVQSPFMTMAAEIKKMGPGA 180

QY 149 TAGGAEKSNVVKIOSTPVKQSGGGCC 173

Db 181 TAGGAEKSNVVKIOSTPVKQSGGGCC 205

RESULT 3

TVMSYP

GTP-binding protein ypt1 - mouse

N/Alternate names: GTP-binding protein Rab1; ras-related protein ypt1; transforming prote

C/Species: Mus musculus (house mouse)

C/Date: 31-Mar-1991 #sequence\_revision 31-Mar-1991 #text\_change 09-Jul-2004

C/Accession: S05551; S06285

R/Wichmann, H.; Disela, C.; Haubruck, H.; Gallwitz, D.

Nucleic Acids Res. 17, 6737-6738, 1989

A/Title: Nucleotide sequence of the mouse ypt1 gene encoding a ras-related GTP-binding p

A/Reference number: S05551; MUID:89386011; PMID:2506528

A/Accession: S05551

A/Molecule type: DNA

A/Residues: 1-205 &lt;MTC&gt;

A/Cross-references: UNIPROT:P11476; EMBL:X15744; NID:g55458; PIDN:CAA33760.1; PID:g76315f

R/Habdruck, R.; Disela, C.; Wagner, P.; Gallwitz, D.

EMBO J. 6, 4049-4053, 1987

A/Title: The ras-related ypt protein is an ubiquitous eukaryotic protein: isolation and e

A/Reference number: S06285; MUID:8816649; PMID:3127202

A/Accession: S06285

A/Molecule type: mRNA

A/Residues: 1-205 &lt;HADU&gt;

C/Genetics:

A/Gene: ypt1

A/Introns: 8/2; 32/3; 64/3; 96/3; 140/3

C/Function:

A/Description: probably involved in protein transport from the endoplasmic reticulum thrc

C/Superfamily: ras transforming protein; translation elongation factor Tu homolog

C/Keywords: endoplasmic reticulum; Golgi apparatus; GTP binding; lipoprotein; membrane p

tein

F:12-127/Domains: translation elongation factor Tu homology &lt;ETU&gt;

F:18-25/Region: nucleotide-binding motif A (P-loop)

F:62-67/Region: nucleotide-binding motif B

F:124-127/Region: GTP-binding NXXD motif

F:154-156/Region: GTP-binding SAK/L motif

F:24,25,43,124,125,127,154/Binding site: Mg-GTP (Lys, Ser, Thr, Asn, Lys, Asp, Ser) #stat

F:194/Binding site: phosphate (Ser) (covalent) (by cdcd kinase) #status predicted

F:204,205/Binding site: geranyl-geranyl (Cys) (covalent) #status predicted

Query Match 97.1%; Score 860; DB 1; Length 205;

Best Local Similarity 84.4%; Pred. No. 1.7e-61;

Matches 173; Conservative 0; Mismatches 0; Indels 32; Gaps 1;

QY 1 MSSNNPEYDYLFLKLLIGDSGVKSCLLRFADDTYESYISTIGVDFKIRITELDGKTI 60

Db 1 MSSNNPEYDYLFLKLLIGDSGVKSCLLRFADDTYESYISTIGVDFKIRITELDGKTI 60

QY 61 KLGQI-----ESNNYKQWLOEIDRYASENVNKL 88

Db 61 KLGQIWDTAGQERFRITTSYRGAGHIIIVYDVTQESFNNVQWLOEIDRYASENVNKL 120

QY 89 LVGNKCDLTTKKVVDYTTAKEPADSLGIPLETSAKATNVQSPFMTMAAEIKKMGPGA 148

Db 121 LVGNKCDLTTKKVVDYTTAKEPADSLGIPLETSAKATNVQSPFMTMAAEIKKMGPGA 180

QY 149 TAGGAEKSNVVKIOSTPVKQSGGGCC 173

Db 181 TAGGAEKSNVVKIOSTPVKQSGGGCC 205

RESULT 4

TVRTYP

GTP-binding protein Rab1 - rat

N/Alternate names: transforming protein ypt1 homolog

C/Species: Rattus norvegicus (Norway rat)

C/Date: 03-Aug-1992 #sequence\_revision 06-Dec-1996 #text\_change 09-Jul-2004

C/Accession: A39963

R/Zahraoui, A.; Touchot, N.; Chardin, P.; Tavittian, A.

Proc. Natl. Acad. Sci. U.S.A. 84, 8210-8214, 1987

A/Title: Four additional members of the ras gene superfamily isolated by an oligonucleot

A/Reference number: A39963; MUID:88068563; PMID:3317403

A/Accession: A39963

A/Molecule type: mRNA

A/Residues: 1-205 &lt;TOU&gt;

A/Cross-references: UNIPROT:P05711; GB:J02998; NID:g206552; PIDN:AAA42006.1; PID:g206553

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: November 17, 2004, 14:08:55 ; Search time 40 Seconds

(without alignments)  
286.825 Million cell updates/sec

Title: US-09-820-003C-2  
Perfect score: 886  
Sequence: 1 MSSMNPBYDYLFTLLIGDS.....EKSNNVKGSTPVKQSGGCC 173

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:  
1: /cgn2\_6/ptodata/1/iaa/5A COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PTUS COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	857	96.7	205	4 US-09-709-103-49	Sequence 49, App1
2	857	96.7	205	4 US-09-439-410A-49	Sequence 49, App1
3	813.5	91.8	201	2 US-08-531-525-13	Sequence 13, App1
4	813.5	91.8	201	2 US-08-718-270A-13	Sequence 13, App1
5	766.5	86.5	201	2 US-08-916-901-3	Sequence 3, App1
6	766.5	86.5	201	2 US-09-154-602-3	Sequence 3, App1
7	760.5	85.8	201	3 US-08-916-901-8	Sequence 8, App1
8	760.5	85.8	201	3 US-09-154-602-8	Sequence 8, App1
9	702.5	79.3	202	2 US-08-531-525-14	Sequence 14, App1
10	702.5	79.3	202	2 US-08-718-270A-14	Sequence 14, App1
11	596.5	67.3	227	4 US-09-248-796A-20291	Sequence 20291, A
12	439	49.5	215	2 US-08-531-525-10	Sequence 10, App1
13	439	49.5	215	2 US-08-718-270A-10	Sequence 10, App1
14	424	47.9	207	2 US-08-531-525-35	Sequence 35, App1
15	424	47.9	207	2 US-08-718-270A-35	Sequence 35, App1
16	418	47.2	207	2 US-08-824-873-4	Sequence 4, App1
17	418	47.2	207	2 US-09-198-184-4	Sequence 4, App1
18	417.5	47.1	205	2 US-08-531-525-25	Sequence 25, App1
19	417.5	47.1	205	2 US-08-718-270A-25	Sequence 25, App1
20	403	45.5	203	4 US-09-255-920A-12	Sequence 12, App1
21	390	44.0	198	2 US-08-531-525-51	Sequence 51, App1
22	390	44.0	198	2 US-08-718-270A-51	Sequence 51, App1
23	366	41.3	194	2 US-08-531-525-34	Sequence 34, App1
24	366	41.3	194	2 US-08-718-270A-34	Sequence 34, App1
25	360	40.6	213	4 US-09-270-767-46812	Sequence 46812, A
26	360	40.6	213	4 US-09-248-796A-20293	Sequence 20293, A
27	334.5	37.8	214	4 US-09-270-767-33012	Sequence 33012, A

28	334.5	37.8	214	4 US-09-270-767-48229	Sequence 48229, A
29	334	37.7	208	2 US-08-531-525-17	Sequence 17, App1
30	334	37.7	208	2 US-08-718-270A-17	Sequence 17, App1
31	333	37.6	212	3 US-09-399-913-67	Sequence 67, App1
32	333	37.6	212	3 US-09-350-614-67	Sequence 67, App1
33	330.5	37.3	218	4 US-08-531-525-19	Sequence 19, App1
34	330.5	37.3	218	4 US-08-718-270A-19	Sequence 19, App1
35	330	37.2	212	2 US-08-531-525-18	Sequence 18, App1
36	330	37.2	212	2 US-08-718-270A-18	Sequence 18, App1
37	325.5	36.7	217	4 US-09-255-920A-15	Sequence 15, App1
38	322	36.3	210	2 US-08-531-525-16	Sequence 16, App1
39	322	36.3	210	2 US-08-718-270A-16	Sequence 16, App1
40	319	36.0	214	2 US-08-531-525-52	Sequence 52, App1
41	319	36.0	214	2 US-08-718-270A-52	Sequence 52, App1
42	317	35.8	128	4 US-09-513-999C-5775	Sequence 5775, App1
43	316.5	35.7	190	2 US-08-824-873-3	Sequence 3, App1
44	316.5	35.7	190	3 US-09-198-184-3	Sequence 3, App1
45	316.5	35.7	190	4 US-09-255-920A-7	Sequence 7, App1

## ALIGNMENTS

```
RESULT 1
US-09-709-103-49
; Sequence 49, Application US/09709103
; Patent No. 6733991
; GENERAL INFORMATION:
; APPLICANT: Cismowski, Mary
; TITLE OF INVENTION: AGS Proteins and Nucleic Acid Molecules and Uses Therefor
; FILE REFERENCES: 60388-A-PC-US
; CURRENT APPLICATION NUMBER: US/09/709,103
; CURRENT FILING DATE: 2000-11-08
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 49
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-709-103-49

Query Match      96.7%; Score 857; DB 4; Length 205;
Best Local Similarity 83.9%; Pred. No. 2, 2e-92;
Matches 172; Conservative 1; Mismatches 0; Indels 32; Gaps 1;

QY 1 MSSMNPBYDYLFTLLIGDSGVKSCLLRFADDTYTESYISTIGVDFKIRTIELDGKTI 60
DB 1 MSSMNPBYDYLFTLLIGDSGVKSCLLRFADDTYTESYISTIGVDFKIRTIELDGKTI 60
QY 61 KLOI-----ESFNNVQWIOEIDRVASEVNVKL 88
DB 61 KLOIWTAGGERFRTTSSYRGAGIIYVYDVTDESFFNNVQWIOEIDRVASEVNVKL 120
QY 89 LVNKKCDLTKKVVDTTAKFADSLGIPIETSAKNAATVVEOSFMTMAEIKRGPGCA 148
DB 121 LVNKKCDLTKKVVDTTAKFADSLGIPIETSAKNAATVVEOSFMTMAEIKRGPGCA 180
QY 149 TAGGAEKSNVKTOSTPVKQSGGCC 173
DB 181 TAGGAEKSNVKTOSTPVKQSGGCC 205

RESULT 2
US-09-439-410A-49
; Sequence 49, Application US/09439410A
; Patent No. 6746852
; GENERAL INFORMATION:
; APPLICANT: Cismowski, Mary
; TITLE OF INVENTION: AGS PROTEIN AND NUCLEOTIC ACID MOLECULES AND USES THEREOF
; FILE REFERENCES: 1919/60388-B
; CURRENT APPLICATION NUMBER: US/09/439,410A
```

CURRENT FILING DATE: 1999-11-11  
 NUMBER OF SEQ ID NOS: 118  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 49  
 LENGTH: 205  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-09-439-410A-49

Query Match 96.7%, Score 857, DB 4, Length 205,  
 Best Local Similarity 83.9%, Pred. No. 2,2e-92,  
 Matches 172; Conservative 1; Mismatches 0; Indels 32; Gaps 1;

QY 1 MSSNNPEYDLFKLLIGDSGVKSCLLLRPADDTYTESYSTIGVDFKRTIELDGKTI 60  
 DB 1 MSSNNPEYDLFKLLIGDSGVKSCLLLRPADDTYTESYSTIGVDFKRTIELDGKTI 60  
 QY 61 KLOI-----ESFNNVKOMLOEIDRYASENVNKL 88  
 DB 61 KLOIWDPAQGERFRITSSYRGAGHIIIVYDVTDQESFNNVKOMLOEIDRYASENVNKL 120  
 QY 89 LVNKKCDLTTKKVVDYTTAKFPADSLGIPLETSAKNAATNVESQSFMTMAEIKRMGPGA 148  
 DB 121 LVNKKCDLTTKKVVDYTTAKFPADSLGIPLETSAKNAATNVESQSFMTMAEIKRMGPGA 180  
 QY 149 TAGGAEKSNVKTQSTPVKQSGGCC 173  
 DB 181 TAGGAEKSNVKTQSTPVKQAGGCC 205

## RESULT 3

US-08-531-525-13  
 Sequence 13, Application US/08531525  
 Patent No. 5840683

GENERAL INFORMATION:  
 APPLICANT: Hlawka, Joseph J.  
 APPLICANT: Pincus, Matthew R.  
 APPLICANT: No. 5840683le, John F.  
 APPLICANT: Abajian, Henry B.  
 APPLICANT: Kende, Andrew S.  
 TITLE OF INVENTION: Peptides Inhibiting the Oncogenic Action  
 TITLE OF INVENTION: of p21 Ras  
 NUMBER OF SEQUENCES: 52  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Greenlee and Winner, P.C.  
 STREET: 5370 Manhattan Circle, Suite 201  
 CITY: Boulder  
 STATE: Colorado  
 COUNTRY: US  
 ZIP: 80303  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/531,525  
 FILING DATE: 21-SEP-1995  
 CLASSIFICATION: 530  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Feiber, Donna M.  
 REGISTRATION NUMBER: 33,878  
 REFERENCE/DOCKET NUMBER: 37-94  
 TELEPHONE: (303) 499-8080  
 TELEFAX: (303) 499-8089  
 INFORMATION FOR SEQ ID NO: 13:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 201 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein

HYPOTHETICAL: NO  
 ORIGINAL SOURCE:  
 ORGANISM: Discopyge ommata  
 US-08-531-525-13

Query Match 91.8%, Score 813.5, DB 2, Length 201,  
 Best Local Similarity 81.2%, Pred. No. 2,7e-87,  
 Matches 164; Conservative 3; Mismatches 2; Indels 33; Gaps 2;

QY 4 MNEPYDLFKLLIGDSGVKSCLLLRPADDTYTESYSTIGVDFKRTIELDGKTIKQ 63  
 DB 1 MNEPYDLFKLLIGDSGVKSCLLLRPADDTYTESYSTIGVDFKRTIELDGKTIKQ 60  
 QY 64 I-----ESFNNVKOMLOEIDRYASENVNKL 91  
 DB 61 IWDPAQGERFRITSSYRGAGHIIIVYDVTDQESFNNVKOMLOEIDRYASENVNKL 120  
 QY 92 NKCDLTTKKVVDYTTAKFPADSLGIPLETSAKNAATNVESQSFMTMAEIKRMGPATAG 151  
 DB 121 NKCDLTTKKVVDYTTAKFPADSLGIPLETSAKNAATNVESQSFMTMAEIKRMGPATAG 179  
 QY 152 GAERKSNVKTQSTPVKQSGGCC 173  
 DB 180 GSEKSNVNIQSTPVKSSGGGCC 201

## RESULT 4

US-08-718-270A-13  
 Sequence 13, Application US/08718270A  
 Patent No. 5910478

GENERAL INFORMATION:  
 APPLICANT: Hlawka, Joseph J.  
 APPLICANT: Pincus, Matthew R.  
 APPLICANT: No. 5910478le, John F.  
 APPLICANT: Abajian, Henry B.  
 APPLICANT: Kende, Andrew S.  
 TITLE OF INVENTION: Peptidomimetics Inhibiting  
 TITLE OF INVENTION: the Oncogenic Action of p21 Ras  
 NUMBER OF SEQUENCES: 52  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Greenlee, Winner and Sullivan, P.C.  
 STREET: 5370 Manhattan Circle, Suite 201  
 CITY: Boulder  
 STATE: Colorado  
 COUNTRY: US  
 ZIP: 80303  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/718,270A  
 FILING DATE: 20-SEP-1996  
 CLASSIFICATION: 514  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/531,525  
 FILING DATE: 21-SEP-1995  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 60/004,091  
 FILING DATE: 21-SEP-1995  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Feiber, Donna M.  
 REGISTRATION NUMBER: 33,878  
 REFERENCE/DOCKET NUMBER: 78-95  
 TELEPHONE: (303) 499-8080  
 TELEFAX: (303) 499-8089  
 INFORMATION FOR SEQ ID NO: 13:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 201 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2004, 13:59:49 ; Search time 156 Seconds  
(without alignments)  
397.822 Million cell updates/sec

Title: US-09-820-003C-2

Perfect score: 886  
Sequence: 1 MSSNRPYDYLFLKLLIGDS.....EKSNVKIQSPVKGSGGCC 173

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729239 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_23Sep04:\*

1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	886	100.0	173	ABG72742	ABG72742 Human Ras
2	886	100.0	173	ADM87199	ADM87199 Human pro
3	860	97.1	205	AAAB34843	AAAB34843 Gene 44 h
4	860	97.1	205	AAAB34844	AAAB34844 Human sec
5	860	97.1	205	ADBB61461	ADBB61461 Amyloid b
6	860	97.1	205	ADBB61416	ADBB61416 Human pro
7	860	97.1	205	ADBB61412	ADBB61412 Human pro
8	859	97.0	205	ADBB61441	ADBB61441 205 amino
9	817	92.2	198	AAAB34816	AAAB34816 Human sec
10	792.5	89.4	169	AAE29159	AAE29159 Human Ras
11	768.5	86.7	221	AAO13525	AAO13525 Human pol
12	768.5	86.7	224	AAAB58758	AAAB58758 Breast an
13	766.5	86.5	201	AAAY00919	AAAY00919 Human Rab
14	766.5	86.5	201	AAU28024	AAU28024 Novel hum
15	766.5	86.5	201	ABU52701	ABU52701 Human int
16	766.5	86.5	201	ADJ76657	ADJ76657 Novel hum
17	760.5	85.8	201	ABU52702	ABU52702 Human hea
18	760.5	85.8	201	ABU52702	ABU52702 Intracell
19	760.5	85.8	201	ADBB63414	ADBB63414 Rat prote
20	760.5	85.8	201	ADBB63410	ADBB63410 Rat prote
21	760.5	85.8	201	ADBB63450	ADBB63450 Rat prote
22	760.5	85.8	201	ADBB63481	ADBB63481 Rat prote
23	760.5	85.8	201	ADBB63400	ADBB63400 Rat prote
24	700	79.0	141	ADAA54719	ADAA54719 Human pro
25	683	77.1	205	ABBS59808	ABBS59808 Drosophila

26	637	71.9	202	3	AAAG07763	AAAG07763 Arabidops
27	635	71.7	225	6	ABJ26183	ABJ26183 Aspergill
28	633	71.4	202	3	AAAG10858	AAAG10858 Arabidops
29	631.5	71.3	203	8	ADG00360	ADG00360 Nicotiana
30	619.5	69.9	258	3	AAAG30498	AAAG30498 Arabidops
31	618.5	69.8	203	3	AAAG30499	AAAG30499 Arabidops
32	612	69.1	218	3	AAAG30710	AAAG30710 Arabidops
33	612	69.1	221	3	AAAG30709	AAAG30709 Arabidops
34	612	69.1	254	3	AAAG45323	AAAG45323 Arabidops
35	596.5	67.3	207	5	ABP73466	ABP73466 Candida a
36	580	65.5	197	3	AAAG48820	AAAG48820 Arabidops
37	571	64.4	206	6	ABRS3078	ABRS3078 Protein s
38	571	64.4	206	7	ADK62970	ADK62970 Disease t
39	513.5	58.0	153	8	ADP29898	ADP29898 Human sec
40	447	50.5	216	8	ADN72851	ADN72851 Thale cre
41	444	50.1	218	3	AAAG19220	AAAG19220 Arabidops
42	442	49.9	216	3	AAAG08006	AAAG08006 Arabidops
43	440	49.7	206	6	ABJ26542	ABJ26542 Aspergill
44	440	49.7	206	6	ABJ25583	ABJ25583 Aspergill
45	431	48.6	215	3	AAAG35215	AAAG35215 Zea mays

#### ALIGNMENTS

RESULT 1	ABG72742	standard; protein; 173 AA.
ID	ABG72742	standard; protein; 173 AA.
XX	ABG72742;	
AC	ABG72742;	
DT	19-FEB-2003	(first entry)
XX	Human Ras-like protein.	
DE	Human Ras-like protein.	
XX	Human, Ras-like protein; Ras; low molecular weight; LMW;	
XX	small regulatory guanine nucleotide-binding protein; GTP-binding protein;	
XX	small G protein; signal transduction; regulation; modulator; diagnosis;	
KW	apoptosis; AIDS; neurodegenerative disease; Alzheimer's disease;	
KW	Parkinson's disease; myelodysplastic syndrome; aplastic anemia;	
KW	ischemic injury; myocardial infarction; stroke; toxin-induced disease;	
KW	alcohol-induced liver damage; cirrhosis; wasting disease; cachexia;	
KW	viral infection; hepatitis; osteoporosis; cell proliferation; cancer;	
KW	leukemia; inflammation; allergy; asthma; atherosclerosis;	
KW	diabetes mellitus; rheumatoid arthritis; therapeutic; transgenic;	
KW	pharmacogenomic analysis; genotype; antisense; gene therapy.	
XX	Homo sapiens.	
OS	US2002142382-A1.	
XX	03-OCT-2002.	
XX	29-MAR-2001; 2001US-00820003.	
XX	29-MAR-2001; 2001US-00820003.	
XX	29-MAR-2001; 2001US-00820003.	
XX	(MERK/) MERKULOV G. V.	
XX	(DFRA/) DI FRANCESCO V.	
XX	(BEAS/) BEASLEY E. M.	
XX	Merkulov GV, Di Francesco V, Beasley EM;	
XX	WPI; 2003-102518/09.	
XX	N-PSDB; ABX13973.	
XX	Novel isolated human Ras-like polypeptide useful for diagnosing,	
XX	preventing and treating inflammation and disorders associated with cell	
XX	proliferation and apoptosis.	
XX	Claim 1; Fig 2; 70pp; English.	
XX	The invention discloses an isolated human Ras-like polypeptide and the	



CC polynucleotide that encodes it. Ras proteins are small (low molecular  
 CC weight, LMW) regulatory guanine nucleotide-binding protein (GTP-binding  
 CC proteins or small G proteins) and are key relays in the signal  
 CC transduction cascade. The Ras-like proteins participate in a wide range  
 CC of regulatory functions. The Ras-like protein is useful for identifying a  
 CC modulator of function or activity and for identifying an agent that binds  
 CC to it. The polypeptide and antibody are useful in the diagnosis,  
 CC prevention and treatment of disorders associated with an increase in  
 CC apoptosis, such as AIDS, neurodegenerative diseases, such as Alzheimer's  
 CC disease, Parkinson's disease, myelodysplastic syndromes, such as aplastic  
 CC anaemia, ischaemic injuries, such as myocardial infarction, stroke, toxin  
 CC -induced diseases, such as alcohol-induced liver damage, cirrhosis,  
 CC wasting diseases, such as cachexia, viral infections, such as hepatitis B  
 CC and C and osteoporosis, cell proliferation disorders, such as cancer and  
 CC leukaemia and inflammation disorders, such as allergy, asthma,  
 CC atherosclerosis, diabetes mellitus and rheumatoid arthritis. The  
 CC polypeptide and polynucleotide are useful as models for the development  
 CC of human therapeutics. The polypeptide is useful to raise antibodies or  
 CC to elicit an immune response, to determine levels of the protein in  
 CC biological fluids and tissues, in drug screening assays and in  
 CC pharmacogenomic analysis. The polynucleotide is useful for constructing  
 CC transgenic animals, for monitoring the effectiveness of modulating  
 CC compounds, for testing an individual for a genotype, as antisense  
 CC constructs, and for gene therapy. The sequence presented is the human Ras-  
 CC like protein

XX Sequence 173 AA;

Query Match 100.0%; Score 886; DB 6; Length 173;  
 Best Local Similarity 100.0%; Pred. No. 5,7e-81;

- Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSNNPEVDYLFKLLIGDSGVGKSCLLRFADPTYSYISTIGVDFKIRTIELDGKTI 60  
 Db 1 MSSNNPEVDYLFKLLIGDSGVGKSCLLRFADPTYSYISTIGVDFKIRTIELDGKTI 60  
 QY 61 KLOIESFNNVQKWLQIEDRYASENVNKKLVGNKCDLITTKVVDYTTAKFADSLGIPFLE 120  
 Db 61 KLOIESFNNVQKWLQIEDRYASENVNKKLVGNKCDLITTKVVDYTTAKFADSLGIPFLE 120  
 QY 121 TSAKATNVEQSFMTMAAEIKKRMGPATAGAGKSNVKIOSTPVKQSGGGCC 173  
 Db 121 TSAKATNVEQSFMTMAAEIKKRMGPATAGAGKSNVKIOSTPVKQSGGGCC 173

RESULT 2

ADM87199

ID ADM87199 standard; protein; 173 AA.

XX ADM87199;

DT 03-JUN-2004 (first entry)

XX Human protein SEQ ID NO:292.

XX respiratory; cytosolic; antiarthritic; antiinflammatory;  
 KM gastrointestinal; antibacterial; immunosuppressive; antidiabetic;  
 KM antirheumatic; gene therapy; molecular weight marker; chromosome marker;  
 KM chromosome tag; genetic fingerprinting; nutritional supplement; cancer;  
 KM inflammatory condition; arthritis; inflammatory bowel disease;  
 KM Crohn's disease; sepsis; rheumatoid arthritis; diabetes mellitus type 1;  
 KM graft versus host disease; human.

XX Homo sapiens.

XX WO2004009834-A2.

XX 29-JAN-2004.

PF 19-JUL-2002; 2002WO-US022858.

PR 21-JUL-2001; 2001US-0306971P.  
 PR 28-MAR-2002; 2002US-00112944.

XX (NUVE-) NUVELO INC.  
 PA  
 XX  
 XX Tang YF, Yang Y, Meng G, Zhang J, Ren F, Xue A, Wang J;  
 PI Weinman T, Ghosh MJ, Wang D, Zhao QH, Wang Z;  
 XX  
 XX WPI; 2004-143291/14.  
 DR N-PSDB; ADM86955.

PT New isolated polynucleotides and polypeptides, useful for treating, e.g.  
 PT cancer, lung or liver fibrosis, arthritis, inflammatory bowel disease,  
 PT Crohn's disease, rheumatoid arthritis, diabetes mellitus type 1 or graft  
 PT versus host disease.

PS Claim 20; SEQ ID NO 292; 591bp; English.

CC The present invention describes an isolated polynucleotide (1): (a)  
 CC comprising a nucleotide sequence selected from SEQ ID NO:1-244; or (b)  
 CC which encodes a polypeptide with biological activity, where the  
 CC polynucleotide hybridizes to (i) under stringent hybridization conditions  
 CC or has greater than 9% sequence identity with (1). (1) has respiratory,  
 CC cytosolic, antiarthritic, antiinflammatory, gastrointestinal,  
 CC antibacterial, immunosuppressive, antidiabetic and antirheumatic  
 CC activities, and can be used in gene therapy. (1) can be used for  
 CC generating polynucleotides encoding chimeric or fusion proteins and  
 CC heterologous protein sequences. The polynucleotides can be used to  
 CC express recombinant protein for analysis, characterization or therapeutic  
 CC use; as markers for tissues in which the corresponding protein is  
 CC preferentially expressed; as molecular weight markers on gels; as  
 CC chromosome markers or tags to identify chromosomes or to map related gene  
 CC positions; to compare with endogenous DNA sequences in patients to  
 CC identify potential genetic disorders; as probes to hybridize and discover  
 CC genes, related DNA sequences; as a source of information to derive PCR  
 CC primers for genetic fingerprinting; as a probe to subtract-out known  
 CC sequences in the process of discovering other novel polynucleotides; for  
 CC selecting and making oligomers for attachment to a gene chip or other  
 CC support, including for examination of expression patterns; and as anti-  
 CC protein antibodies using DNA immunisation techniques; and as an antigen  
 CC to raise anti-DNA antibodies or elicit another immune response. The  
 CC polynucleotides and polypeptides can also be used as nutritional sources  
 CC or supplements, e.g. as a protein or amino acid supplement, as a carbon  
 CC source, as a nitrogen source or as a source of carbohydrates. The  
 CC polynucleotides and polypeptides can also be used treat cancer. The  
 CC compositions are useful for promoting better or faster closure of non-  
 CC healing wounds, for the generation and regeneration of tissues, for gut  
 CC protection or regeneration and treatment of lung or liver fibrosis,  
 CC reperfusion injury in various tissues, and conditions resulting from  
 CC systemic cytokine damage. The compositions can also be used to treat  
 CC inflammatory conditions (e.g. arthritis, inflammatory bowel disease or  
 CC Crohn's disease), sepsis, rheumatoid arthritis, diabetes mellitus type 1  
 CC or graft versus host disease. The present sequence represents a novel  
 CC human polypeptide sequence from the present invention. N.B. The sequences  
 CC for this patent were obtained from the USPTO web site from an equivalent  
 CC US patent US20040048249A1.

XX Sequence 173 AA;

Query Match 100.0%; Score 886; DB 6; Length 173;  
 Best Local Similarity 100.0%; Pred. No. 5,7e-81;

- Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSNNPEVDYLFKLLIGDSGVGKSCLLRFADPTYSYISTIGVDFKIRTIELDGKTI 60  
 Db 1 MSSNNPEVDYLFKLLIGDSGVGKSCLLRFADPTYSYISTIGVDFKIRTIELDGKTI 60  
 QY 61 KLOIESFNNVQKWLQIEDRYASENVNKKLVGNKCDLITTKVVDYTTAKFADSLGIPFLE 120  
 Db 61 KLOIESFNNVQKWLQIEDRYASENVNKKLVGNKCDLITTKVVDYTTAKFADSLGIPFLE 120  
 QY 121 TSAKATNVEQSFMTMAAEIKKRMGPATAGAGKSNVKIOSTPVKQSGGGCC 173  
 Db 121 TSAKATNVEQSFMTMAAEIKKRMGPATAGAGKSNVKIOSTPVKQSGGGCC 173





QY	60	CGTTTGCTCCCGGAAACAGCTATCTCATCTCTTTCTTTGATTAACCGTGGCCGAG	119
Db	72	CGTTTGCTCCCGGAAACAGCTATCTCATCTCTTTCTTTGATTAACCGTGGCCGAG	131
QY	120	AGTCAGGCGCGCGCTCGCGCAGCAGAGGGCGCGGTGGCGCGCGCAGCTGAC	179
Db	132	AGTCAGGCGCGCGCTCGCGCAGCAGAGGGCGCGGTGGCGCGCGCAGCTGAC	191
QY	180	ATTGCCAGCATGAAATCCCGAAATAGTATTATTTCAAGTTACTCTGATTTGGACTCA	239
Db	192	ATTGCCAGCATGAAATCCCGAAATAGTATTATTTCAAGTTACTCTGATTTGGACTCA	251
QY	240	GGGGTTGAAAGCTTGCTCTCTCTTGAAGTTTGACATATACATACAGAAAGCTAC	299
Db	252	GGGGTTGAAAGCTTGCTCTCTCTTGAAGTTTGACATATATATACAGAAAGCTAC	311
QY	300	ATACAGCAATTTGGTGGATTTTCAAAATAGAACTATAGAGTTAGACGGGAAACAAATC	359
Db	312	ATACAGCAATTTGGTGGATTTTCAAAATAGAACTATAGAGTTAGACGGGAAACAAATC	371
QY	360	AAAGCTTCAATA-----	371
Db	372	AAAGCTTCAATA-----	431
QY	372	-----GAGTCTTCAAT	383
Db	432	TAACAGAGAGCCATGGCATATAGTTGGTATGATGTGACAGATCAGGAGTCTTCAAT	491
QY	384	AAATGTTAAACAGGGCTGCAGAGAAATAGATCGTATCCGACAGGAAATGTCAAAATTTG	443
Db	492	AAATGTTAAACAGGGCTGCAGAGAAATAGATCGTATCCGACAGGAAATGTCAAAATTTG	551
QY	444	TTGGTAGGGAAACAATGTATCTGACACAAAGAAAGTAGTAGATACACAACAGCGAAG	503
Db	552	TTGGTAGGGAAACAATGTATCTGACACAAAGAAAGTAGTAGATACACAACAGCGAAG	611
QY	504	GAATTTGCGATTCCTTGGAAATCCGTTTTTGGAAACAGGTCTAAGATGCAAGAAAT	563
Db	612	GAATTTGCGATTCCTTGGAAATCCGTTTTTGGAAACAGGTCTAAGATGCAAGAAAT	671
QY	564	GTAGAAACAGTCTTTCATGACGATGAGCTGAGATTAAAGCGAATGGGTCCCGGAGA	623
Db	672	GTAGAAACAGTCTTTCATGACGATGAGCTGAGATTAAAGCGAATGGGTCCCGGAGA	731
QY	624	ACAGCTGGTGGTGGTGAAGTCCAAATGTTAAATTCAGAGCACTTCCAGTCAAGCAGTCA	683
Db	732	ACAGCTGGTGGTGGTGAAGTCCAAATGTTAAATTCAGAGCACTTCCAGTCAAGCAGTCA	791
QY	684	GGTGAAGGTGTGCTGAATAATTTGGCTCATCTTTTCTCAAGCAATGAATTTGCATTC	743
Db	792	GGTGAAGGTGTGCTGAATAATTTGGCTCATCTTTTCTCAAGCAATGAATTTGCATTC	851
QY	744	TGAACCCAAAGTAAAAAACAATAATTCCTGAATTTACTGTATGTAGCTGACTACAACA	803
Db	852	TGAACCCAAAGTAAAAAACAATAATTCCTGAATTTACTGTATGTAGCTGACTACAACA	911
QY	804	GATTCCTAACCGCTCCACAAAGGTGAGATTGTAATAGTCAATATGACACTTTTTTTTTT	863
Db	912	GATTCCTAACCGCTCCACAAAGGTGAGATTGTAATAGTCAATATGACACTTTTTTTTTT	971
QY	864	ATTCCCTTGAAGTCAAGACAGCTAATCTCAATTTTGAAGCTGTTTAAACCTTTGTGTCT	923
Db	972	ATTCCCTTGAAGTCAAGACAGCTAATCTCAATTTTGAAGCTGTTTAAACCTTTGTGTCT	1031
QY	924	GGTTATATAAATATAGTGTGTAATCCTTGTGCTTTCTGATACCAAGCTGTTTCCGCTG	983
Db	1032	GGTTATATAAATATAGTGTGTAATCCTTGTGCTTTCTGATACCAAGCTGTTTCCGCTG	1091
QY	984	GTGGTTGAATATATTTGTTTGTATGTTTATATGTGCAAGTTTAAATGTCAGGTTTAA	1043
Db	1092	GTGGTTGAATATATTTGTTTGTATGTTTATATGTGCAAGTTTAAATGTCAGGTTTAA	1151

QY	1044	TCCTCTGAAGATGAAGTCACCACTTTGTATGCAAAACGCAACAAGCAGTGTCTGCACCTT	1103
Db	1152	TCCTTCTGAAGATGAAGTTCAGCCATTTTGTATGCAAAACGCAACAAGCAGTGTCTGCACCTT	1211
QY	1104	TCCATGACATTAAGTTTACGATGATTTATATGATAGACCTGATTTGCTAGTCTTCCCTTG	1163
Db	1212	TCCATGACATTAAGTTTACGATGATTTATATGATAGACCTGATTTGCTAGTCTTCCCTTG	1271
QY	1164	TAGAGTTTAATTAATGAAGAATTACACATCTCGATTAATAGTTTCTTCATCTCTGCACATAT	1223
Db	1272	TAGAGTTTAATTAATGAAGAATTACACATCTCGATTAATAGTTTCTTCATCTCTGCACATAT	1331
QY	1224	AATTTGTGCTGCAGAAATATTGTAAATTTGTGTGCACACTATGTAACAAAACAACATGAAGAT	1283
Db	1332	AATTTGTGCTGCAGAAATATTGTAAATTTGTGTGCACACTATGTAACAAAACAACATGAAGAT	1391
QY	1284	ATGTTTAATTAATTAATTTGACTTATTTGGAAGTAAAAAAA	1323
Db	1392	ATGTTTAATTAATTAATTTGACTTATTTGGAAGTAAAAAAA	1431

RESULT 2				
CR593530	CR593530	1440 bp	mRNA	linear
LOCUS	full-length cDNA clone CSODC023YH6 of Neuroblastoma Cot			
DEFINITION	25-normalized of Homo sapiens (human).			
ACCESSION	CR593530			
VERSION	CR593530.1			
KEYWORDS	HTC; CNSLT_CDNA.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			

REFERENCE	AUTHORS	TITLE	JOURNAL	REMARK
1 (bases 1 to 1440)	Li, W. B., Gruber, C., Jesse, J. and Polayes, D.	Full-length cDNA libraries and normalization	Unpublished	Contact : Feng Liang Email : <a href="mailto:liang@lifetech.com">liang@lifetech.com</a> URL : <a href="http://lifetech.com">lifetech.com</a>

REFERENCE 2 (bases 1 to 1440)  
AUTHORS Genoscope.  
TITLE Direct Submission  
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :

**COMMENT**  
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalised. Library was constructed by Life Technologies, a division of Invitrogen.

```
FEATURES
source
location/Qualifiers
1..1440
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSDDC023YH06"
/tissue_type="Neuroblastoma Cot 25-normalized"
/plasmid="pCMVSPORT_6"
```

Query Match	84.6%	Score 1188;	DB 3;	Length 1440;
Best Local Similarity	93.1%	Pred. No. 4.9e-196;		
Matches 1305; Conservative	0;	Mismatches 0;	Indels 97;	Gaps 2;

QY	Db	QY
1 AAGCATTAGCTAGG- GGGGGGGGCTGCTGATTTGGTTCTAGGGGAGCGAGTAGGGGAAAGA	39 AAGCATAGCTAGTACGCGCGCGCGCTGCTGATTTGGTTCTAGGGGAGCGAGTAGGGGAAAGA	59
60 CGTTGGTCTCCCGGAAAGCCTTATCTCATTCCTTTCTTTCGATTACCGTGGGCGGAG		98
99 CGTTGGTCTCCCGGAAAGCCTTATCTCATTCCTTTCTTTCGATTACCGTGGGCGGAG		158



QY	181	TEGTCAGCAATGAATCCCGAATATGATTAATTTATTCAAAGTTACTTCGATTTGGCGACTCAG	240
Db	181	TGTCACGCAATGAATCCCGAATATGATTAATTTATTCAAAGTTACTTCGATTTGGCGACTCAG	240
QY	241	GGGTTGGAAAGTCCTTGCCCTTCCTCTTAGTGGTTGGCAGATGATACATATACAGAAAGCTACA	3000
Db	241	GGGTTGGAAAGTCCTTGCCCTTCCTCTTAGTGGTTGGCAGATGATACATATACAGAAAGCTACA	3000
QY	301	TCAGCACAATTGGTGTGATATTTCAAATAAGAACTATAGATTAGACGGGAAAACAATCA	3600
Db	301	TCAGCACAATTGGTGTGATATTTCAAATAAGAACTATAGATTAGACGGGAAAACAATCA	3600
QY	361	AGCTTCAAAATAGAGTCCTTCATATATGTTTAAACAGGGCGCGCAGGAAATGATCGTTATG	4200
Db	361	AGCTTCAAAATAGAGTCCTTCATATATGTTTAAACAGGGCGCGCAGGAAATGATCGTTATG	4200
QY	421	CCAGTAAAAATGTCAACAAATTTGGTAGGGAACAAATGTGATCTGACCACAAACAAAG	4800
Db	421	CCAGTAAAAATGTCAACAAATTTGGTAGGGAACAAATGTGATCTGACCACAAACAAAG	4800
QY	481	TGTTAGACTACCAACAGGGAAGAGATTTGCTGATCCCTGGAAATTCGGTTTTTGGAAA	5400
Db	481	TGTTAGACTACCAACAGGGAAGAGATTTGCTGATCCCTGGAAATTCGGTTTTTGGAAA	5400
QY	541	CCAGTCTAAGATATGCACAGCAAGATGTAGAAACAGCTTTCATGACGATGCGAGCTGATTA	6000
Db	541	CCAGTCTAAGATATGCACAGCAAGATGTAGAAACAGCTTTCATGACGATGCGAGCTGATTA	6000
QY	601	AAAAAGGATGGGTCCCGGAGCAACAGCTGTGTGTGTCTGAAAGTCCAAATGTTAAAAATTC	6600
Db	601	AAAAAGGATGGGTCCCGGAGCAACAGCTGTGTGTGTCTGAAAGTCCAAATGTTAAAAATTC	6600
QY	661	AGAGCACTCCAGTCAAGCAGTCAGGTGAGAGTTGCTGCTAAATAATTCCTCCATCCTTTT	7200
Db	661	AGAGCACTCCAGTCAAGCAGTCAGGTGAGAGTTGCTGCTAAATAATTCCTCCATCCTTTT	7200
QY	721	CTCACAAGCAATGAATTTGGCAATCTGAACCCAAAGTAAAAACAAATAATTCCTGAAATTTGA	7800
Db	721	CTCACAAGCAATGAATTTGGCAATCTGAACCCAAAGTAAAAACAAATAATTCCTGAAATTTGA	7800
QY	781	CTGTATGTAGTCGCACTACAAACAGATTCTTACCGTCTCCACAAAGGTACAGAAATTTGTA	8400
Db	781	CTGTATGTAGTCGCACTACAAACAGATTCTTACCGTCTCCACAAAGGTACAGAAATTTGTA	8400
QY	841	TGTCCAATACCTGATTTTTTTTTTAAATTCCTTGAACCTCAAGACAGCTTACATTTTCAGA	9000
Db	841	TGTCCAATACCTGATTTTTTTTTTAAATTCCTTGAACCTCAAGACAGCTTACATTTTCAGA	9000
QY	901	ACTGTTTAAACCTTGTGTGTCTGATTTTAAATAATGTGTGTATCCTTGTGCTTTC	9600
Db	901	ACTGTTTAAACCTTGTGTGTCTGATTTTAAATAATGTGTGTATCCTTGTGCTTTC	9600
QY	961	CTGATACAGACTGTTTCCCGTGTGTGATTAATATTTTGTGTTGATGTTTATTTG	10200
Db	961	CTGATACAGACTGTTTCCCGTGTGTGATTAATATTTTGTGTTGATGTTTATTTG	10200
QY	1021	GCATGTTAAATGTACAGGTTTAATCTTCTGAAAGTGAAGTTCAAGCAATTTGTATCAAC	10800
Db	1021	GCATGTTAAATGTACAGGTTTAATCTTCTGAAAGTGAAGTTCAAGCAATTTGTATCAAC	10800
QY	1081	AGCAACAAGCACTGTCTGTCACTTTCACAGCAATTAAGTTATGAGATCTTATATGTAGA	11400
Db	1081	AGCAACAAGCACTGTCTGTCACTTTCACAGCAATTAAGTTATGAGATCTTATATGTAGA	11400
QY	1141	TCTGATTTGCTAGTTCTTCTGTGATTAATTAATGGAAGATTACACTATCTGATTA	12000
Db	1141	TCTGATTTGCTAGTTCTTCTGTGATTAATTAATGGAAGATTACACTATCTGATTA	12000
QY	1201	TAGTTTCTTCACTCTGCAATATAATTTGGCTGCGAGAAATTTGTATATTTGTTGCACAC	12600
Db	1201	TAGTTTCTTCACTCTGCAATATAATTTGGCTGCGAGAAATTTGTATATTTGTTGCACAC	12600

QY	1261	TAGTAAACAAAACACATGGAATAGTTTAATAAATATGTAATCTATTGGAAGTAAGAAA	1320
Db	1261	TATGTACAAAACCAACTGAAGATAGTTTAATAAATAATTTGTAATCTATTGGAAGTAAGAAA	1320
QY	1321	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	1380
Db	1321	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	1380
QY	1381	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	1405
Db	1381	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	1405
RESULT 2			
US-10-252-157-411			
; Sequence 411, Application US/10252157			
; Publication No. US20030190640A1			
; GENERAL INFORMATION:			
; APPLICANT: Farib, Mary			
; TITLE OF INVENTION: GENES EXPRESSED IN PROSTATE CANCER			
; FILE REFERENCE: PA-0027-1 US			
; CURRENT APPLICATION NUMBER: US/10/252,157			
; PRIOR FILING DATE: 2002-10-01			
; PRIOR APPLICATION NUMBER: 60/295,048			
; NUMBER OF SEQ ID NOS: 501			
; SOFTWARE: PERL Program			
; SEQ ID NO 411			
; LENGTH: 8131			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: misc_feature			
; OTHER INFORMATION: Inycle ID No. US20030190640A1 411474.17			
; FEATURE:			
; NAME/KEY: unsure			
; LOCATION: 3486-3786			
; OTHER INFORMATION: a, t, c, g, or other			
US-10-252-157-411			
Query Match 85.5%; Score 1201.2; DB 15; Length 8131;			
Best Local Similarity 93.0%; Pred. No. 9e-193;			
Matches 120; Conservative 0; Mismatches 3; Indels 97; Gaps 2;			
QY	1	AAGCATAGCTGAGT-GCGGCGCTGCTGATGTGTGTTCTAAGGGACGAGTAGGGAGA	59
Db	21	AAGGATAGCTGAGTGCAGCGCGCTGCTGATGTGTGTTAAGGGACGAGTAGGGAGA	80
QY	60	CGTTGCTCTCCCGGACAGCCTATCTCATTCCTTCTTTCGATTCACCGTGGCGCGAG	119
Db	81	CGTTGCTCTCCCGGACAGCCTATCTCATTCCTTCTTTCGATTCACCGTGGCGCGAG	140
QY	120	AGTAGAGCGCGCGCTGCGGACCAAGGCGCGGTGGCGCGCGCGCGCGAGCTGCAGTGC	179
Db	141	AGTAGAGCGCGCGCTGCGGACCAAGGCGCGGTGGCGCGCGCGCGCGAGCTGCAGTGC	200
QY	180	ATGTCCAGCATGAATCCCGAATATGATATTTATTCGAATTAACCTTGATTTGGCGACTCA	239
Db	201	ATGTCCAGCATGAATCCCGAATATGATATTTATTCGAATTAACCTTGATTTGGCGACTCA	260
QY	240	GGGGTTGAAAAGTCTTGCCCTTCTTCTTAAGTTTGCAGATATACATACAGAAAAGCTAC	299
Db	261	GGGGTTGAAAAGTCTTGCCCTTCTTCTTAAGTTTGCAGATATACATACAGAAAAGCTAC	320
QY	300	ATACGACCAATTGGTGTGATTTCAAAAATAAGACTATAGATTAGACGGGAAAACAATC	359
Db	321	ATACGACCAATTGGTGTGATTTCAAAAATAAGACTATAGATTAGACGGGAAAACAATC	380
QY	360	AAGCTTCAATA-----	371
Db	381	AAGTTCAATATAGGACACAGTAGGCTACGAATAATTTGATACATAATACCTTCACTTAT	440



QY	298	ACATTCAGCAAAATGGGTGTGATTTTCAAAATTAAGAACTATAGAGTTAAGACGGGAAAAACA	357
Db	321	ACATTCAGCAAAATGGGTGTGATTTTCAAAATTAAGAACTATAGAGTTAAGACGGGAAAAACA	380
QY	358	TCAGACTTCAAAAT-----	371
Db	381	TCAAAGCTTCAAATATATGAGACACAGACAGGCCAGAAAGATTTGAAACAATCACCTCAGTT	440
QY	372	-----GAGTCTTTCA	381
Db	441	ATTACAGAGAGGCCCATGGCATATAGTGTGTATGATGTGACAGATCAGAGAGTCTCTTCA	500
QY	382	ATAATGTTAAACAGTGGCTTCGAGAAATAGATGCTTATGCCAGTGAATAATGTCAACAAT	441
Db	501	ATAATGTTAAACAGTGGCTTCGAGAAATAGATGCTTATGCCAGTGAATAATGTCAACAAT	560
QY	442	TGTTGGTAGGAAACAAATGTGATCTGACCACAAAGAAATAGTAGACTACACACAGCA	501
Db	561	TGTTGGTAGGAAACAAATGTGATCTGACCACAAAGAAATAGTAGACTACACACAGCA	620
QY	502	AGGAATTTGCTGATTCCTTGGAAATCCGTTTGGAAACCAAGTGTAGAATGACAGA	561
Db	621	AGGAATTTGCTGATTCCTTGGAAATCCGTTTGGAAACCAAGTGTAGAATGACAGA	680
QY	562	ATGTAGAACAGTCTTTCATGACGATGGCAGCTGAGATTAAAAAGCGATGGGTCCCGAG	621
Db	681	ATGTAGAACAGTCTTTCATGACGATGGCAGCTGAGATTAAAAAGCGATGGGTCCCGAG	740
QY	622	CAACAGCTGCTGGTGTGAGAAAGTCCAAAGTTTAAATTCAGAGCACTCCAGTCAAGCAGT	681
Bb	741	CAACAGCTGCTGGTGTGAGAAAGTCCAAAGTTTAAATTCAGAGCACTCCAGTCAAGCAGT	800
QY	682	CAGGTGAGAGTTCGCTGTAATAATTTGCTCCATCCCTTTTCTCACAGCATGAATTTGCA	741
Db	801	CAGGTGAGAGTTCGCTGTAATAATTTGCTCCATCCCTTTTCTCACAGCATGAATTTGCA	860
QY	742	TCTGAACCCCAAGTGAACAAATTTGCTGCAATTTGTAAGTATGTAAGTGTGCACTACAA	801
Db	861	TCTGAACCCCAAGTGAACAAATTTGCTGCAATTTGTAAGTATGTAAGTGTGCACTACAA	920
QY	802	CAGATTTCTTACCGTCTCCACAAGGTTCAGAGATTTGAATGGTTCATCTGACTTTT	861
Db	921	CAGATTTCTTACCGTCTCCACAAGGTTCAGAGATTTGAATGGTTCATCTGACTTTT	980
QY	862	TTATTCCTTGACTCAACACAGCTACTTTCACAGACTGTTTAAACCTTGTGTG	921
Db	981	TTATTCCTTGACTCAACACAGCTACTTTCACAGACTGTTTAAACCTTGTGTG	1040
QY	922	CTGGTTTAAATAATGATGTGTATTCCTGTGTGCTTCCTGATACCAAGACTGTTCCG	981
Db	1041	CTGGTTTAAATAATGATGTGTATTCCTGTGTGCTTCCTGATACCAAGACTGTTCCG	1100
QY	982	TGGTTGGTTAGAAATATTTGTTTGAATGTTTATATTTGGAAGTGTTAAGATGTCAAGTTT	1041
Db	1101	TGGTTGGTTAGAAATATTTGTTTGAATGTTTATATTTGGAAGTGTTAAGATGTCAAGTTT	1160
QY	1042	AGTCTTCGAAGATAGAGTTCAGCAATTTGATATCAACACAGACAGAGGTCTGTGAC	1101
Db	1161	AGTCTTCGAAGATAGAGTTCAGCAATTTGATATCAACACAGACAGAGGTCTGTGAC	1220
QY	1102	TTTCCATSCATAAAGTTTATAGTAGATGTTATATGTAAGATCTGAATTTGCTAGTTCTTCT	1161
Db	1221	TTTCCATSCATAAAGTTTATAGTAGATGTTATATGTAAGATCTGAATTTGCTAGTTCTTCT	1280
QY	1162	TGTAGAGTTATTAATGGAAGATTACATATCTGATTATATGTTTCTTCATATCTGTGAT	1221
Db	1281	TGTAGAGTTATTAATGGAAGATTACATATCTGATTATATGTTTCTTCATATCTGTGAT	1340
QY	1222	ATAATTTTGTGCTGAGAAATATGTAAATTTGTGAGACATACTGTAAACAAACACAGAG	1281
Db	1341	ATAATTTTGTGCTGAGAAATATGTAAATTTGTGAGACATACTGTAAACAAACACAGAG	1400
QY	1282	ATAATGTTTAAATATATTTACTTATTTGGAAGTAAAAAAA	1323

```

Db      1401 ATATGTTTATAAATATGTGACTTATTGTGAAGTAATATCAA 1442

RESULT 2
US-09-016-434-1422
/ Sequence 1422, Application US/0901434
/ Patent No. 6500938
/ GENERAL INFORMATION:
/ APPLICANT: Janice Au-Young
/ APPLICANT: Jeffrey U. Seihamer
/ TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
/ TITLE OF INVENTION: PATHWAY GENE EXPRESSION
/ NUMBER OF SEQUENCES: 1490
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
/ STREET: 3174 PORTER DRIVE
/ CITY: PALO ALTO
/ STATE: CALIFORNIA
/ COUNTRY: USA
/ ZIP: 94304
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/016,434
/ FILING DATE: HERewith
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER:
/ FILING DATE:
/ CLASSIFICATION:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Zeller, Karen J.
/ REGISTRATION NUMBER: 37,071
/ REFERENCE/DOCKET NUMBER: PA-0002 US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (650) 855-0555
/ TELEFAX: (650) 845-4166
/ INFORMATION FOR SEQ ID NO: 1422:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 723 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ LIBRARY: GENBANK
/ CLONE: G550059
/
US-09-016-434-1422

Query Match      33.9%; Score 476.2; DB 4; Length 723;
Best Local Similarity 84.6%; Pred. No. 7,4e-88;
Matches 613; Conservative 0; Mismatches 13; Indels 99; Gaps 3;

QY      129 GGGCGGCTGGCGGCACCAAGCGCGGGTGGCGCGCGCGGAGCTTCAGTGCATATCTCCAGC 188
Db      2   GGGCGCGGTGGCGCAGCAAGCGCGGGTG--CGCGCGCAGCTGCAGTGCATGTCCAGC 59
QY      189 ATGAATCCCGAATATGATTAATTAATTAATCAAGTTACTTCTGATTTGGCGACTCGAGGGTTGGA 248
Db      60 ATGAATCCCGAATATGATTAATTAATTAATCAAGTTACTTCTGATTTGGCGACTCGAGGGTTGGA 119
QY      249 AAGTCTTGCTCTTCTTTAGGTTTGCAGATGATATCATATACAGAAAGCTTACATCAGACA 308
Db      120 AAGCTTGCTCTTCTTTAGGTTTGCAGATGATATCATATACAGAAAGCTTACATCAGACA 179
QY      309 ATTGGTGTGATTTCAAAATTAAGACATATAGAGTTATACGGGAAAACAATCAAGCTTCAA 368
Db      180 ATTGGTGTGATTTCAAAATTAAGACTATATAGAGTTATACGGGAAAACAATCAAGCTTCAA 239
QY      369 ATA----- 371

```





DR P-PSDB; ABG72742.

XX Novel isolated human Ras-like polypeptide useful for diagnosing,  
PT preventing and creating inflammation and disorders associated with cell  
PT proliferation and apoptosis.

XX Claim 4; Fig 1; 70pp; English.

XX The invention discloses an isolated human Ras-like polypeptide and the  
CC polynucleotide that encodes it. Ras proteins are small (low molecular  
CC weight, LMW) regulatory guanine nucleotide-binding protein (GTP-binding  
CC proteins or small G proteins) and are key relays in the signal  
CC transduction cascade. The Ras-like proteins participate in a wide range  
CC of regulatory functions. The Ras-like protein is useful for identifying a  
CC modulator of function or activity and for identifying an agent that binds  
CC to it. The polypeptide and antibody are useful in the diagnosis,  
CC prevention and treatment of disorders associated with an increase in  
CC apoptosis, such as AIDS, neurodegenerative diseases, such as Alzheimer's  
CC disease, Parkinson's disease, myelodysplastic syndromes, such as aplastic  
CC anaemia, ischaemic injuries, such as myocardial infarction, stroke, toxin  
CC -induced diseases, such as alcohol-induced liver damage, cirrhosis,  
CC and C and osteoporosis, cell proliferation disorders, such as hepatitis B  
CC leukemia and inflammation disorders, such as allergy, asthma,  
CC atherosclerosis, diabetes mellitus and rheumatoid arthritis. The  
CC polypeptide and polynucleotide are useful as models for the development  
CC of human therapeutics. The polypeptide is useful to raise antibodies or  
CC biological fluids and tissues, in drug screening assays and in  
CC pharmacogenomic analysis. The polynucleotide is useful for constructing  
CC transgenic animals, for monitoring the effectiveness of modulating  
CC compounds, for testing an individual for a genotype, as antisense  
CC constructs and for gene therapy. The sequence presented is the human Ras-  
CC like protein cDNA

XX Sequence 1405 BP; 471 A; 234 C; 300 G; 400 T; 0 U; 0 Other;

Query Match 100.0%; Score 1405; DB 10; Length 1405;  
Best Local Similarity 100.0%; Pred. No. 2e-183;  
Matches 1405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 AAGCATAGCTAGAGCGGCGGCGGCTGATGTTCTTCAAGGCGGCGAGTNGGGGAAGAC 60
DB 1 AAGCATAGCTAGAGCGGCGGCGGCTGATGTTCTTCAAGGCGGCGAGTNGGGGAAGAC 60
QY 61 GTTGTCTCTCCCGAAGACGCTATCTCAATCTCTTCTTCAATTAACCGTGGCGGAGA 120
DB 61 GTTGTCTCTCCCGAAGACGCTATCTCAATCTCTTCTTCAATTAACCGTGGCGGAGA 120
QY 121 GTCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
DB 121 GTCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
QY 181 TGTTCAGCAGTAATCCGAATATGATTAATTAATTAATTAATTAATTAATTAATTAATTA 240
DB 181 TGTTCAGCAGTAATCCGAATATGATTAATTAATTAATTAATTAATTAATTAATTAATTA 240
QY 241 GGGTTGGAAGAGTCTTGGCTCTTCTTCAAGTTCGATGATGATGATGATGATGATGATGAT 300
DB 241 GGGTTGGAAGAGTCTTGGCTCTTCTTCAAGTTCGATGATGATGATGATGATGATGATGAT 300
QY 301 TCAGCACAATGTTGTGATTTCAAAAATAGAACTATAGATTAGCGGGAAGAAACATCA 360
DB 301 TCAGCACAATGTTGTGATTTCAAAAATAGAACTATAGATTAGCGGGAAGAAACATCA 360
QY 361 AGCTTCAATAGAGTCTTCAATTAATGTTAAACAGTGTGAGGAAATAGATCGTTATG 420
DB 361 AGCTTCAATAGAGTCTTCAATTAATGTTAAACAGTGTGAGGAAATAGATCGTTATG 420
QY 421 CCAGTGAATATGTCACAAATTTGTTGGTGGGAAACAAATGATATGACCAACAAAGAAAG 480
DB 421 CCAGTGAATATGTCACAAATTTGTTGGTGGGAAACAAATGATATGACCAACAAAGAAAG 480

```

```

QY 481 TAGTACTACACACAGCGAAGGAAATTTGCTGATTCCTTGGAATTCGTTTGGAAA 540
DB 481 TAGTACTACACACAGCGAAGGAAATTTGCTGATTCCTTGGAATTCGTTTGGAAA 540
QY 541 CCAGTGTCAAGAAATGCAACAAATGTAAGACAGTCTTTCATGAGATGAGGAGTGAATTA 600
DB 541 CCAGTGTCAAGAAATGCAACAAATGTAAGACAGTCTTTCATGAGATGAGGAGTGAATTA 600
QY 601 AAAGCGAAATGAGTCCCGGAGCAACAGCTGTGTGTCTGAGAAATGTTAAATTC 660
DB 601 AAAGCGAAATGAGTCCCGGAGCAACAGCTGTGTGTCTGAGAAATGTTAAATTC 660
QY 661 AGAGCACTCAGTCAAGCAAGTCAAGTGTGAGTGTGCTGTAATAATTCCTCATCTTTT 720
DB 661 AGAGCACTCAGTCAAGCAAGTCAAGTGTGAGTGTGCTGTAATAATTCCTCATCTTTT 720
QY 721 CTACAGCAATGAAATTTGCAATCTGAACCAAGTGAACAAATTTGCTGAATTTGA 780
DB 721 CTACAGCAATGAAATTTGCAATCTGAACCAAGTGAACAAATTTGCTGAATTTGA 780
QY 781 CTGATGATGCTGCACTACACAGATTTCTTACCGTCCCAACAGTCAAGATTTGAAA 840
DB 781 CTGATGATGCTGCACTACACAGATTTCTTACCGTCCCAACAGTCAAGATTTGAAA 840
QY 841 TGGTCAATCTGACTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 900
DB 841 TGGTCAATCTGACTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 900
QY 901 ACGTTTAAACCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
DB 901 ACGTTTAAACCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
QY 961 CTGATACAGACTGTTTCCCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1020
DB 961 CTGATACAGACTGTTTCCCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1020
QY 1021 GGATGTTTGAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1080
DB 1021 GGATGTTTGAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1080
QY 1081 AGCACAAGCAGTGTGCTGCACTTCCATGCAATTAAGTTTGTGAGTGTGAGTGTGAGTGT 1140
DB 1081 AGCACAAGCAGTGTGCTGCACTTCCATGCAATTAAGTTTGTGAGTGTGAGTGTGAGTGT 1140
QY 1141 TGTGATTTGCTGATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1200
DB 1141 TGTGATTTGCTGATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1200
QY 1201 TAGTTTCTTCAATCTGCAATTAATTTGTGCTGCAATTAATTTGTGCTGCAATTAATTT 1260
DB 1201 TAGTTTCTTCAATCTGCAATTAATTTGTGCTGCAATTAATTTGTGCTGCAATTAATTT 1260
QY 1261 TATGTAAACAAACCACTGAAGATATGTTTAAATTAATTAATTAATTAATTAATTAATTA 1320
DB 1261 TATGTAAACAAACCACTGAAGATATGTTTAAATTAATTAATTAATTAATTAATTAATTA 1320
QY 1321 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1380
DB 1321 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1380
QY 1381 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1405
DB 1381 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1405

```

# RESULT 2

ADP03021 standard; cDNA; 1433 BP.

ADP03021;

29-JUL-2004 (first entry)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 18, 2004, 09:13:54 ; Search time 6027 Seconds  
(without alignments)  
11024.063 Million cell updates/sec

Title: US-09-820-003C-1

Perfect score: 1405

Sequence: 1 aagcgatagctgagtcgagcgc.....aaaaaaaaaaaaaaaaaaaaa 1405

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:  
1: gb\_ha:\*  
2: gb\_hcg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sbs:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vt:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1277	90.9	1625	9 HSM806623	BX571747 Homo sapi
2	1219	86.8	1433	9 BC000905	BC000905 Homo sapi
3	1179.2	83.9	8137	6 AR454534	AR454534 Sequence
4	1103.6	78.5	2257	6 AX713964	AX713964 Sequence
5	1103.6	78.5	2257	9 AK055927	AK055927 Homo sapi
6	1052.6	74.9	1444	10 BC002077	BC002077 Mus muscu
7	1044	74.3	1208	9 HSM800023	AL050268 Homo sapi
8	1036.2	73.8	2819	10 AK129477	AK129477 Mus muscu
9	1028.2	73.2	1425	10 BC066662	BC066662 Rattus no
10	1013.2	72.1	1428	10 MMYPT1	Y00094 Mouse mRna
11	955	68.0	2686	10 AF226873	AF226873 Mus muscu
12	820.6	58.4	292895	2 AC123143	AC123143 Rattus no
13	817.2	58.2	204230	9 AC007318	AC007318 Homo sapi
14	797.4	56.8	329753	2 AC096701	AC096701 Rattus no
15	777.8	55.4	327906	2 AC117841	AC117841 Rattus no
16	741.4	52.8	203071	10 AL606522	AL606522 Mouse DNA
17	727.8	51.8	243263	2 AC132510	AC132510 Rattus no
18	697.2	49.6	327906	2 AC117841	AC117841 Rattus no
19	687.6	48.9	116614	2 AL512349	AL512349 Homo sapi

c 20	686	48.8	104081	9 HSDJ26337	AL049545 Human DNA
c 21	684.4	48.7	162471	9 AC067945	AC067945 Homo sapi
c 22	682.6	48.6	4678	10 MMYPT14	X15747 Mouse YP1
c 23	674	48.0	217009	2 AC094972	AC094972 Rattus no
c 24	651	46.3	703	11 BV166639	BV166639 RAB1A_104
c 25	622.6	44.3	1006	11 BV105532	BV105532 MARC_2615
c 26	611.6	43.5	191930	2 AC120240	AC120240 Rattus no
c 27	573.8	40.8	310641	2 AC099354	AC099354 Rattus no
c 28	573.2	40.8	840	4 CFRAB1	X56384 Canine rab1
c 29	567.6	40.4	141036	2 AC023807	AC023807 Mus muscu
c 30	543.4	38.7	1086	10 RN063023	U63023 Rattus norv
c 31	506	36.0	506	6 AX396159	AX396159 Sequence
c 32	476.2	33.9	723	6 AR270859	AR270859 Sequence
c 33	476.2	33.9	723	9 HUMRAB1A	M28209 Homo sapien
c 34	467	33.2	487	6 CQ050923	CQ050923 Sequence
c 35	467	33.2	487	6 CQ065951	CQ065951 Sequence
c 36	467	33.2	487	6 CQ092963	CQ092963 Sequence
c 37	467	33.2	487	6 CQ131774	CQ131774 Sequence
c 38	467	33.2	487	6 CQ170346	CQ170346 Sequence
c 39	467	33.2	487	6 CQ199514	CQ199514 Sequence
c 40	467	33.2	487	6 CQ214966	CQ214966 Sequence
c 41	467	33.2	487	6 CQ253556	CQ253556 Sequence
c 42	467	33.2	487	6 CQ290658	CQ290658 Sequence
c 43	467	33.2	487	6 CQ327680	CQ327680 Sequence
c 44	455.8	32.4	110000	2 AC096343	AC096343 Rattus no
c 45	452.2	32.2	110000	2 AC094411_1	Continuation (2 of

#### ALIGNMENTS

RESULT 1  
HSM806623  
LOCUS  
DEFINITION  
HSM806623 Homo sapiens mRNA; cDNA DKFZp686E08159 (from clone DKFZp686E08159); complete cds.  
ACCESSION  
BX571747  
KEYWORDS  
BX571747.1 GI:33096734  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
Koehler,K., Beyer,A., Mewes,H.W., Weil,B., Amlid,C., Osanger,A., Fobbo,G., Han,W. and Wiemann,S.  
The German Human cDNA Consortium  
Direct Submission  
Submitted (16-JUL-2003) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY  
COMMENT  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
Heinrich-Heine-University, Dueseldorf(Germany) within the cDNA sequencing consortium of the German Genome Project. This clone (DKFZp686E08159) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heudenweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at <http://mips.gsf.de/proj/cDNA/>.  
FEATURES  
source  
1..1625  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/map="2p14"  
/clone="DKFZp686E08159"  
/tissue\_type="human colon endothel primary cell culture"  
/clone\_lib="686 (synonym: hlcc3). Vector pSPORT1\_Sfi; host DH10B; sites SfiIA + SfiIB  
/dev\_stage="adult"  
1..1625  
/gene="DKFZp686E08159"

## CDS

184..801  
 /gene="DKFZp686E08159"  
 /note="RAB1A, member RAS oncogene family"  
 /codon\_start=1  
 /product="hypothetical protein"  
 /protein\_id="CAE11872.1"  
 /db\_xref="GI:3306735"  
 /translation="MSSMNPEDYLFLKLLIGDSGVKSCLLRFADDTYESYIT  
 KMDLPIRINASENVNKLVGKCDLTKKVDYATYKAEFADSLGIPLETSAKMN  
 VEOSFTMAAIEIKRMGPAATAGAKSNVXIQSTPVKQSGGGCC"  
 1391..1396  
 /gene="DKFZp686E08159"  
 1411  
 /gene="DKFZp686E08159"  
 ORIGIN

Query Match 90.9%; Score 1277; DB 9; Length 1625;  
 Best Local Similarity 93.5%; Pred. No. 1e-194;  
 Matches 1405; Conservative 0; Mismatches 0; Indels 98; Gaps 3;

```

QY 1 AAGGATAGCTGAGT-GGGCGGCTGCTGATTTGTTCTTGGGACGAGTAGGGAGA 59
Db 4 AAGGATAGCTGAGTGGCGGCTGCTGATTTGTTCTTGGGACGAGTAGGGAGA 63
QY 60 GGTTCCTCTCCGGAAACAGCTATCTCATTCCTTCTTTCGATTACCCGTGGCGGAG 119
Db 64 GGTTCCTCTCCGGAAACAGCTATCTCATTCCTTCTTTCGATTACCCGTGGCGGAG 123
QY 120 AGTCAGGCGGCGGCTGCGGACAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGAG 179
Db 124 AGTCAGGCGGCGGCTGCGGACAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGAG 183
QY 180 ATGTCCAGCATGATCCGAATATGATTTATTTCAAGTTACTCTGATGGGAGCTCA 239
Db 184 ATGTCCAGCATGATCCGAATATGATTTATTTCAAGTTACTCTGATGGGAGCTCA 243
QY 240 GGGGTTGAAAAGCTTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 299
Db 244 GGGGTTGAAAAGCTTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 303
QY 300 ATCAGCAATTTGGTGTGATTTCAAAATAGAACTATAGATTAGACGGGAAAACATC 359
Db 304 ATCAGCAATTTGGTGTGATTTCAAAATAGAACTATAGATTAGACGGGAAAACATC 363
QY 360 AAGCTTCAATA----- 371
Db 364 AAGCTTCAATAATGAGGACACAGCGCCAGGAAAGATTGCAACATCACCTCCAGTTAT 423
QY 372 -----GAGTCCCTTCAT 383
Db 424 TACAGAGAGCCCATGCGATCATGTTGTGATGATGTGACAGATCAGGAGTCTTCAT 483
QY 384 AATGTTAAACAGTGGCTGCGAGGAAATGATCGTTATGCCAGTAAATGTCACAAATTG 443
Db 484 AATGTTAAACAGTGGCTGCGAGGAAATGATCGTTATGCCAGTAAATGTCACAAATTG 543
QY 444 TTGCTAGGAAACAATGTGATGTGACCAAGAAAGTGTAGACTACACACAGGAG 503
Db 544 TTGCTAGGAAACAATGTGATGTGACCAAGAAAGTGTAGACTACACACAGGAG 603
QY 504 GAATTTGCTGATTCCTTGGAAATTCGTTTGGAAACAGTGTAAAGATGCAACGAT 563
Db 604 GAATTTGCTGATTCCTTGGAAATTCGTTTGGAAACAGTGTAAAGATGCAACGAT 663
QY 664 GTTGAACAGCTTTCATGACGATGAGCTGAGATTAAACCGATGTCGCGGAGCA 723
Db 664 GTTGAACAGCTTTCATGACGATGAGCTGAGATTAAACCGATGTCGCGGAGCA 723
QY 624 ACAGCTGTGTGCTGAGAGTGTCAATGTTAAATTCAGAGCACTTCAGTCAAGAGTCA 683
Db 724 ACAGCTGTGTGCTGAGAGTGTCAATGTTAAATTCAGAGCACTTCAGTCAAGAGTCA 783

```

QY 684 GGTGAGGTTGCTGTAAAAATTTGGCTTCATCTTTCTACAGCAATGAAATTTGCATC 743  
 Db 784 GGTGAGGTTGCTGTAAAAATTTGGCTTCATCTTTCTACAGCAATGAAATTTGCATC 843  
 QY 744 TGAACCAAGTGAACAAACAAATTTGCTGATTTGATCTATGATGCTGACATCAACA 803  
 Db 844 TGAACCAAGTGAACAAACAAATTTGCTGATTTGATCTATGATGCTGACATCAACA 903  
 QY 804 GATTTTACCGTCTCCACAAAGGTCAGAGATTTGAATGTCATATCTGAC-TTTTTTTT 862  
 Db 904 GATTTTACCGTCTCCACAAAGGTCAGAGATTTGAATGTCATATCTGAC-TTTTTTTT 963  
 QY 863 TATTCCTGACTCAACACGCTTACTTCAATTTTCAGAACTGTTTAAACCTTTGTGTC 922  
 Db 964 TATTCCTGACTCAACACGCTTACTTCAATTTTCAGAACTGTTTAAACCTTTGTGTC 1023  
 QY 923 TGTATTAAATTAATGTGTGTAATCTTGTGCTTCTTCTGATACAGACTGTTCCGT 982  
 Db 1024 TGTATTAAATTAATGTGTGTAATCTTGTGCTTCTTCTGATACAGACTGTTCCGT 1083  
 QY 983 GGTTCGTAATATATTTTGTGTGATGTTTATATTTGATGATTTAGATGCAAGTTA 1042  
 Db 1084 GGTTCGTAATATATTTTGTGTGATGTTTATATTTGATGATTTAGATGCAAGTTA 1143  
 QY 1043 GTCTTCGAAGATGAAGTTCAGCAATTTTGTATCAACAGCAACGAGTCTGTCACT 1102  
 Db 1144 GTCTTCGAAGATGAAGTTCAGCAATTTTGTATCAACAGCAACGAGTCTGTCACT 1203  
 QY 1103 TTCCAGCATTAAGTTAGTAGATGTTATATGTAAGATCTGATTTGCTAGTCTCTT 1162  
 Db 1204 TTCCAGCATTAAGTTAGTAGATGTTATATGTAAGATCTGATTTGCTAGTCTCTT 1263  
 QY 1163 GTAGAGTTAATAATGAAAGATTAACATCTGATTTAATGTTCTTCATCTGTGATA 1222  
 Db 1264 GTAGAGTTAATAATGAAAGATTAACATCTGATTTAATGTTCTTCATCTGTGATA 1323  
 QY 1223 TAATTTGGCTGCAAGATTTGTAATTTGTCACACTATGTAAACAACAACCTGAAG 1282  
 Db 1324 TAATTTGGCTGCAAGATTTGTAATTTGTCACACTATGTAAACAACAACCTGAAG 1383  
 QY 1283 TATGTTAATAATTAATGTAATTTGTAATTTGTCACACTATGTAAACAACAACCTGA 1342  
 Db 1384 TATGTTAATAATTAATGTAATTTGTAATTTGTCACACTATGTAAACAACAACCTGA 1443  
 QY 1343 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1402  
 Db 1444 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1503  
 QY 1403 AAA 1405  
 Db 1504 AAA 1506

## REFERENCE

1 (bases 1 to 1433)  
 Klausner,R.D., Collins,F.S., Wagner,L.H., Derge,J.G.,  
 Altschul,S.F., Zeeberg,B., Bueltow,K.H., Schaefer,C.F., Bhat,N.K.,  
 Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,U., Hsieh,F.,  
 Diatchenko,L., Maruine,K., Farmer,A.A., Rubin,G.M., Hong,L.,  
 Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,  
 Scheetz,T.E., Brownstein,M.J., Urdin,T.B., Toshiyuki,S.,

## SOURCE

Homo sapiens (human)  
 BUKARYOTA; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

## LOCUS

BC000905 1433 bp mRNA linear PRI 29-UN-2004  
 MGC:5233 IMAGE:290705), complete cds.

ACCESSION BC000905  
 VERSION BC000905.2 GI:14705268  
 KEYWORDS MGC.

DEFINITION Homo sapiens RAB1A, member RAS oncogene family, mRNA (cDNA clone  
 IMAGE:290705), complete cds.

FEATURES  
 CDS  
 ORIGIN

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: November 17, 2004, 14:17:35 ; Search time 140 Seconds  
(without alignments)  
437.601 Million cell updates/sec

Title: US-09-820-003C-2

Perfect score: 886

Sequence: 1 MSSNMPDYDLFKLLIGDS.....EKSXVXIQTSPVKGSGGCC 173

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1570615 seqs, 354127592 residues

Total number of hits satisfying chosen parameters: 1570615

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBSCOMB.pep:\*

2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*

3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*

4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBSCOMB.pep:\*

5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*

6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBSCOMB.pep:\*

7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*

8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBSCOMB.pep:\*

9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBSCOMB.pep:\*

10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBSCOMB.pep:\*

11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBSCOMB.pep:\*

12: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*

13: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBSCOMB.pep:\*

14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBSCOMB.pep:\*

15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBSCOMB.pep:\*

16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBSCOMB.pep:\*

17: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*

18: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep:\*

19: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*

20: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBSCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	886	100.0	173	9	US-09-820-003A-2
2	886	100.0	173	15	US-10-112-944-292
3	860	97.1	222	9	US-09-820-003A-4
4	857	96.7	205	16	US-10-804-491-49
5	812.5	91.7	201	17	US-10-471-411-6
6	792.5	89.4	169	17	US-10-471-411-2
7	768.5	86.7	224	9	US-09-928-298-466
8	768.5	86.7	224	14	US-10-102-806-466
9	766.5	86.5	201	9	US-09-967-736-3
10	766.5	86.5	201	14	US-10-291-172-193
11	766.5	86.5	201	15	US-10-221-278-193
12	766.5	86.5	201	16	US-10-408-765A-1596
13	766.5	86.5	201	17	US-10-471-411-4

14	760.5	85.8	201	9	US-09-967-736-8	Sequence 8, Appl1
15	759.5	85.7	201	17	US-10-471-411-5	Sequence 5, Appl1
16	700	79.0	141	14	US-10-094-749-2287	Sequence 2287, Appl
17	678	76.5	204	17	US-10-471-411-7	Sequence 7, Appl1
18	660	74.5	205	14	US-10-369-493-6261	Sequence 6261, Ap
19	658.5	74.3	203	17	US-10-471-411-8	Sequence 8, Appl1
20	645	72.8	202	15	US-10-424-599-196551	Sequence 196551,
21	635	71.7	225	14	US-10-128-714-8241	Sequence 8241, Ap
22	624	70.4	225	15	US-10-425-115-47782	Sequence 47782, A
23	624	70.4	239	15	US-10-425-114-41359	Sequence 41359, A
24	624	70.4	240	15	US-10-425-114-42447	Sequence 42447, A
25	624	70.4	240	15	US-10-425-114-56916	Sequence 56916, A
26	623	70.3	202	15	US-10-424-599-186767	Sequence 186767,
27	622	70.2	203	16	US-10-425-115-362779	Sequence 362779,
28	621.5	70.1	203	17	US-10-767-701-445295	Sequence 445295, A
29	621.5	70.1	203	17	US-10-425-115-318489	Sequence 318489,
30	621.5	70.1	225	15	US-10-425-114-47497	Sequence 47497, A
31	621.5	70.1	225	15	US-10-425-114-60367	Sequence 60367, A
32	621.5	70.1	225	15	US-10-425-114-67150	Sequence 67150, A
33	620.5	70.0	213	15	US-10-424-599-186764	Sequence 186764,
34	617	69.6	203	16	US-10-767-701-44811	Sequence 44811, A
35	616.5	69.6	202	16	US-10-437-963-123354	Sequence 123354,
36	615.5	69.5	203	17	US-10-425-115-318493	Sequence 318493,
37	615.5	69.5	228	15	US-10-425-114-60163	Sequence 60163, A
38	615.5	69.5	228	15	US-10-425-114-62895	Sequence 62895, A
39	607.5	68.6	205	15	US-10-425-114-61060	Sequence 61060, A
40	605	68.6	204	17	US-10-425-115-318496	Sequence 318496,
41	599	67.6	216	15	US-10-424-599-186768	Sequence 186768,
42	596.5	67.3	207	14	US-10-032-585-7303	Sequence 7303, Ap
43	591.5	66.8	212	17	US-10-425-115-200936	Sequence 200936,
44	589	66.5	304	16	US-10-437-963-136472	Sequence 136472,
45	578.5	65.3	205	17	US-10-425-115-331973	Sequence 331973,

## ALIGNMENTS

RESULT 1

US-09-820-003A-2

Sequence 2, Application US/09820003A

Patent No. US20020142382A1

GENERAL INFORMATION:

APPLICANT: MERCK/COV, Genmady et al.

TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE

TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE

TITLE OF INVENTION: PROTEINS, AND USES THEREOF

FILE REFERENCE: CLO01196

CURRENT APPLICATION NUMBER: US/09/820,003A

CURRENT FILING DATE: 2001-03-29

NUMBER OF SEQ ID NOS: 39

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 2

LENGTH: 173

TYPE: PRT

ORGANISM: Homo sapien

US-09-820-003A-2

Query Match 100.0%; Score 886; DB 9; Length 173;  
Best Local Similarity 100.0%; Pred. No. 3.9e-78;  
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MSSNMPDYDLFKLLIGDSGVKSCLLAFADDTYESTYSTGTGVDFFKRTLELDKTI	60
Db	1	MSSNMPDYDLFKLLIGDSGVKSCLLAFADDTYESTYSTGTGVDFFKRTLELDKTI	60
Qy	61	KLQIESNNYKQWLOEIDRYASENVNKLAVGNKCDLTTKKVVDYTTAKEFADSLGIPFLE	120
Db	61	KLQIESNNYKQWLOEIDRYASENVNKLAVGNKCDLTTKKVVDYTTAKEFADSLGIPFLE	120
Qy	121	TSANATNVQSFWTMAAEIKKMGPGATNGAEKSNVKIQSTPVGKSGGCC	173
Db	121	TSANATNVQSFWTMAAEIKKMGPGATNGAEKSNVKIQSTPVGKSGGCC	173

```
RESULT 2
US-10-112-944-292
; Sequence 292, Application US/10112944
; Publication No. US20040048249A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Yang, Yonghong
; APPLICANT: Meng, Gezhi
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Xue, Aidong J.
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Wehrman, Tom
; APPLICANT: Ghosh, Malabika
; APPLICANT: Wang, Dunrui
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Zhiwei
; TITLE OF INVENTION: No. US20040048249A1el Nucleic Acids and
; FILE REFERENCE: 805A
; CURRENT APPLICATION NUMBER: US/10/112,944
; CURRENT FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: US 09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: US 09/496,914
; PRIOR FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: US 09/515,126
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: US 09/519,705
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: US 09/552,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: US 09/577,408
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 924
; SOFTWARE: Pf_Fl_genes Version 5.0
; SEQ ID NO 292
; LENGTH: 173
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-112-944-292

Query Match          100.0%; Score 886; DB 15; Length 173;
Best Local Similarity 100.0%; Pred. No. 3,9e-78;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSNNPEYDYLFKLLIGDSGVGKSCLLRFADDTYTESYSTIGVDFKIRITELDGKTI 60
DB 1 MSSNNPEYDYLFKLLIGDSGVGKSCLLRFADDTYTESYSTIGVDFKIRITELDGKTI 60
QY 61 KLOIESNNYKQWQOEIDRYASENVNKLIVGNKCDLTTKKVVDYTTAKEFADSLGIPIELE 120
DB 61 KLOIESNNYKQWQOEIDRYASENVNKLIVGNKCDLTTKKVVDYTTAKEFADSLGIPIELE 120
QY 121 TSAKNATNVBOSFMTMAAEIKKRMGPATAGAEKSNVKIQTPTVKOSGGGCC 173
DB 121 TSAKNATNVBOSFMTMAAEIKKRMGPATAGAEKSNVKIQTPTVKOSGGGCC 173

RESULT 3
US-09-820-003A-4
; Sequence 4, Application US/09820003A
; Patent No. US20020142382A1
; GENERAL INFORMATION:
; APPLICANT: MERKULOV, Gennady et al.
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
```

```
; FILE REFERENCE: CL001196
; CURRENT APPLICATION NUMBER: US/09/820,003A
; CURRENT FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-820-003A-4

Query Match          97.1%; Score 860; DB 9; Length 222;
Best Local Similarity 84.4%; Pred. No. 1,8e-75;
Matches 173; Conservative 0; Mismatches 0; Indels 32; Gaps 1;

QY 1 MSSNNPEYDYLFKLLIGDSGVGKSCLLRFADDTYTESYSTIGVDFKIRITELDGKTI 60
DB 18 MSSNNPEYDYLFKLLIGDSGVGKSCLLRFADDTYTESYSTIGVDFKIRITELDGKTI 77
QY 61 KLOI-----ESFNNVQWLOEIDRYASENVNKL 88
DB 78 KLOIWDTAGQERFRITSSYRGANGIIVYVDYDQESFNNVQWLOEIDRYASENVNKL 137
QY 89 LVGNKCDLTTKKVVDYTTAKEFADSLGIPIELTSAKNATNVBOSFMTMAAEIKKRMGPGA 148
DB 138 LVGNKCDLTTKKVVDYTTAKEFADSLGIPIELTSAKNATNVBOSFMTMAAEIKKRMGPGA 197
QY 149 TAGAEKSNVKIQTPTVKOSGGGCC 173
DB 198 TAGAEKSNVKIQTPTVKOSGGGCC 222

RESULT 4
US-10-804-491-49
; Sequence 49, Application US/10804491
; Publication No. US20040180375A1
; GENERAL INFORMATION:
; APPLICANT: Cismowski, Mary
; APPLICANT: Duzic, Emil
; TITLE OF INVENTION: AGS Proteins and Nucleic Acid Molecules and Uses Therefor
; FILE REFERENCE: 60388-A-PCT-US
; CURRENT APPLICATION NUMBER: US/10/804,491
; CURRENT FILING DATE: 2004-03-19
; PRIOR APPLICATION NUMBER: US/09/709,103
; PRIOR FILING DATE: 2000-11-08
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 49
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-804-491-49

Query Match          96.7%; Score 857; DB 16; Length 205;
Best Local Similarity 83.9%; Pred. No. 3,3e-75;
Matches 172; Conservative 1; Mismatches 0; Indels 32; Gaps 1;

QY 1 MSSNNPEYDYLFKLLIGDSGVGKSCLLRFADDTYTESYSTIGVDFKIRITELDGKTI 60
DB 1 MSSNNPEYDYLFKLLIGDSGVGKSCLLRFADDTYTESYSTIGVDFKIRITELDGKTI 60
QY 61 KLOI-----ESFNNVQWLOEIDRYASENVNKL 88
DB 61 KLOIWDTAGQERFRITSSYRGANGIIVYVDYDQESFNNVQWLOEIDRYASENVNKL 120
QY 89 LVGNKCDLTTKKVVDYTTAKEFADSLGIPIELTSAKNATNVBOSFMTMAAEIKKRMGPGA 148
DB 121 LVGNKCDLTTKKVVDYTTAKEFADSLGIPIELTSAKNATNVBOSFMTMAAEIKKRMGPGA 180
QY 149 TAGAEKSNVKIQTPTVKOSGGGCC 173
DB 181 TAGAEKSNVKIQTPTVKOSGGGCC 205
```